METHODS FOR DEVELOPING AND ASSESSING CIRCUIT MODELS FOR BIPOLAR DIODES AND TRANSISTORS

BY
MENG-KAI CHEN

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1985

ACKNOWLEDGMENTS

I would like to express my gratitude to the chairman of my supervisory committee, Dr. Fredrik A. Lindholm, for his invaluable guidance and assistance during the course of this study. His expertise in semiconductor device physics and outstanding tutoring helped activate my research studies. It is my privilege and my pleasure to be his student. I thank Professor Robert M. Fox for his guidance in the circuit-application aspects of this work. Special thanks go to Mr. Thomas Ifstrom for conducting measurements on lateral p/n/p transistors and Mark Gose for studying the stability of a voltage regulator IC. Thanks are extended to Professors Jerry G. Fossum, Arnost Neugroschel, William R. Eisenstadt, and Tim J. Andersen for their participation on my supervisory committee.

I am grateful to my colleagues and friends, Dr. Tam Wen Jung, Dr. Konstantinos Misiakos, Surya Vasuraghavan, Honggeun Jeong, and Kenjesin Wu, for helpful discussions. Encouragement from Wesley Vincent, Dr. Paul Ainslie, and Richard Entringer is greatly appreciated.

I am indebted to my family for their love and encouragement.

The financial support of the Semiconductor Research Corporation is acknowledged.

TABLE OF CONTENTS

		Page
ACKNOWLEDGMENTS		ii
LIST OF SYMBOLS		vii
ABSTRACT		ix
CHAPTER		
ONE	INTRODUCTION	1
TWO	REVIEW OF EARLIER WORK ON TRANSIENT AND AC MODELING OF THE QUASI-NEUTRAL REGIONS	6
2 1	List of General Equations and the Approximations that Commonly Are Used for Bipolar Semiconductor Devices	6
2 2	Physics of the Delay Phenomenon in P/N Junction Devices	11
2 3	Related Earlier Research	14
2 3 1	Bipolar Junction Diodes	14
2 3 2	Bipolar Junction Transistors	16
THREE	IMPROVED CALCULATIONS OF DELAY AND THEIR APPLICATION TO BIPOLAR JUNCTION DIODES	18
3 1	Area Approach	20
3 2	Ramp Approach	23
3 3	Applications	26
3 3 1	Delay in the Convection Current	28
3 3 2	Delay in the Storage Current	34
3 4	Generalization to Large Signal Excitations	40
3 5	Summary and Discussion	40
FOUR	SYSTEMATIC NON-QUASI-STATIC SMALL-SIGNAL MODELING FOR SEMICONDUCTOR JUNCTION TRANSISTORS AND DIODES	43
4 1	Systematic Inclusion of Delay for Transistors and Diodes	44
4 2	Conversion among Different Configurations	52
4 3	Small-Signal Assessment	55
4 3 1	Comparison with Exact Solution	55

4 3 2 Comparison with Experiments
4 4 Summary 61

FIVE SYSTEMATIC INCLUSION OF DETAILS OF RESPONSE IN LARGE-SIGNAL BIPOLAR TRANSISTOR MODELS THE LARGE-SIGNAL L AND NC MODELS 63

5 1 Introduction 64
5 1 1 Lynn et al 64
5 1 2 The Two-Port Approach 68
5 2 Simplification of the Frequency Dependence of the Transistor Models 70
5 2 1 Truncation for the Vertical Branch of the π-Network 72
5 2 2 Truncation for the Horizontal Branch of the π Network 73
5 3 The Large-Signal NC Model 77
5 4 The Large Signal L Model 80
5 4 1 Options for Deriving the Large Signal L Model 80
5 4 2 Derivation of the Large-Signal L Model 81
5 5 Comparison of the Large-Signal L and NC Models 85
5 5 1 Initial Condition I(0) 86
5 5 2 Computational Efficiency 87
5 6 Small-Signal NC and L Models 88
5 7 Higher-Order Approximation 89
5 7 1 Seitchik's Model 90
5 7 2 Higher-Level Approximations 91
5 8 Extensions of the L and NC Models 93
5 9 Discussion 95
5 9 1 Taylor Expansion About Z = 0 101
5 9 2 Padé Expansions 99
5 10 Summary 102

SIX COMPARISON AND EXTENSION OF RECENT ONE-DIMENSIONAL BIPOLAR TRANSISTOR MODELS 104

6 1 Introduction 104
6 2 Brief Sketch of Non-Quasi-Static Models 105
6 2 1 Partitioned Charge-Based Model (PCB) 105
6 2 2 Transient-Integral-Charge-Control Relation (TICC) 107
6 2 3 Negative Capacitance (NC) and Inductive (L) Equivalent Circuits 108
6 3 Comparison of PCB TICC and NC Models 111
6 3 1 Large-Signal Comparison 111
6 3 2 Small-Signal Comparison 115
6 4 L Model 116
6 4 1 Small-Signal Comparison 116
6 4 2 Large-Signal Comparison 118
6 5 PCB TICC and NC Models Possible Fundamental Limitations 120
6 5 1 Fundamental Limitation of PCB TICC and NC Models 120
6 5 2 Quantitative Explanations 120
6 5 3 Solutions to the Limitation of PCB TICC and NC

Models 121
6 6 Summary and Discussion 123

SEVEN COMPUTER IMPLEMENTATION OF THE LARGE-SIGNAL L AND NC MODELS USING USER DEFINED CONTROL SOURCES 125

7 1 Introduction 125
7 2 Numerical Integration and Nonlinear Solution Algorithms 126
7 2 1 Nonlinear Solution Algorithm 126
7 2 2 Numerical Integration Algorithm 127
7 3 Implementation of the NC Model 129
7 3 1 Implementation of i_{QS} 129
7 3 2 Implementation of $dq_{C(NC)}/dt$ Using the Implicit Trapezoidal Method Oscillation 130
7 3 3 Implementation of $dq_{C(NC)}/dt$ Using the Gear-2 Method 135
7 4 Implementation of the L Model 136
7 4 1 Implementation of $dq_{C(L)}/dt$ Using the Gear-2 Method 137
7 4 1 1 Overflow problem 139
7 4 1 2 Programming subtleties 140
7 4 2 Implementation of $dq_{C(L)}/dt$ Using the Implicit Trapezoidal Method 141
7 4 3 Comparison of the Implicit Trapezoidal Method and the Gear-2 Method 144
7 5 Discussion 145
7 6 Application A Simple Transistor Inverter 146
7 7 Summary 149

EIGHT A SELF-CONSISTENCY TEST FOR DEVICE MODELS IN TRANSIENT COMPUTER SIMULATION OF ELECTRONIC CIRCUITS 153

8 1 Introduction 153
8 2 Self-Consistency 155
8 2 1 Self Consistency from Extended Ideal Charge-Control Model for the Intrinsic Device 155
8 2 2 Model for the Extrinsic Device 167
8 2 3 Self Consistency Test for Large Signal Excitations 170
8 2 4 Self-Consistency Test for Small Signal Sinusoidal Excitations 171
8 3 Model Updating 171
8 4 BJT Inverter Case Study 174
8 5 Discussion 183
8 6 Summary 184

NINE THE IMPORTANCE OF THE NON QUASI-STATIC MODEL FOR CIRCUIT APPLICATIONS 186

9 1 Low-Frequency Applications Lateral p/n/p Bipolar Transistors 186
9 1 1 Case Study A Series Voltage Regulator 186

9 1 1 Level Shifting in an Operational Amplifier 203
9 2 High Frequency Small Signal Application 208
9 3 Summary 212

TEN SUMMARY AND RECOMMENDATIONS 213

APPENDIX

A MATHEMATICAL DETAILS OF DERIVATION OF THE LARGE-SIGNAL L MODEL 215

B UDCS/SLICE/SPICE MODEL IMPLEMENTATION 218

C TRANSFORMATION OF THE NC AND L MODELS TO CHARGE-CONTROL AND PCB REPRESENTATIONS 223

D DETERMINATION OF ELEMENTS OF TICC AND PCB MODELS FOR AN EXPONENTIALLY DOPED BASE 226

REFERENCES 230

BIOGRAPHICAL SKETCH 236

LIST OF SYMBOLS

A	device area
C_{qe}	capacitance (small-signal) derived by quasi-static approach (C/V)
Δx	width of subregion (cm)
$D_{n,p}$	electron, hole diffusivity (cm^2/s)
$\vec{E}$	electric field
$e(t)$	step response
$e'(t)$	the derivative of step response (impulse response)
ε	dielectric constant
η	field factor
$FC(x)$	weight function of the change of the base charge with respect to the collector
$FE(x)$	weight function of the change of the base charge with respect to the emitter
g_m	transconductance (A/V)
G_{qe}	conductance (small signal) derived by quasi static approach (A/V)
i_{QS}	total time-varying electron current
$\vec{J}_{p,\,n}$	hole, electron current density
kT/q	thermal voltage
L^*	complex diffusion length (cm)
L_n	electron diffusion length
n_E	excess electron density at the emitter edge of quasi-neutral base

n_i n_{ie}	intrinsic electron density effective intrinsic [illegible] heavily doped semiconductors
N_{AA}	acceptor impurity concentration
n_{QS}	quasi static excess electron density
$\mu_{p\ n}$	hole electron mobility
q	magnitude of electronic charge (C)
q_{QS}	total excess electron charge in the quasi-neutral base layer
q_C	electron charge related to the collector
q_E	electron charge related to the emitter
s	the complex frequency variable ($\sigma+j\omega$)
t	time
v_{BE}	total time varying base emitter voltage
y_{21}	small signal transconductance
τ_{21}	time constant defined by Eq. (6.5)
τ_{tr}	quasi-static transit time ($X_B^2/2D_n$ for a uniformly doped base)
x	position
X_B	thickness of quasi-neutral base layer (cm)
ω	angular frequency
$\tau_{D(c)}$	delay time associated with the convection current
$\tau_{D(s)}$	delay time associated with the storage current current

Abstract of Dissertation Presented to the Graduate School of the University of Florida in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

METHODS FOR DEVELOPING AND ASSESSING CIRCUIT MODELS FOR BIPOLAR DIODES AND TRANSISTORS

BY

MENG KAI CHEN

August 1989

Chairman: Fredrik A. Lindholm
Major Department: Electrical Engineering

This dissertation focuses on methods for developing and assessing large- and small-signal models for diodes and transistors. Nearly all existing models rely on the quasi-static approximation, which neglects the distributed nature of devices. New non-quasi-static (NQS) models are derived from the Shockley equations that are compact enough for use in circuit simulators. The new negative capacitance (NC) and inductive (L) models are assessed by comparison with experiment, with exact analytical solutions when possible, and with the most widely used quasi-static model, the Gummel and Poon model in SPICE2.

I also compare the new NC and L models, derived by truncation in the complex frequency domain, against other recently proposed NQS

models the partitioned charge based (PCB) model and the transient integral charge control (TICC) model. For forward active operation and an exponentially doped base, the NC model is equivalent to the TICC model. For a uniformly doped base the NC, TICC and PCB models are equivalent. These three models however predict that the magnitude of the transconductance rises with frequency for high frequencies which is incorrect. Truncation in the complex-frequency domain provides a better model the L model.

The newly developed models are implemented in circuit simulators by employing user-defined-controlled-source (UDCS) subroutines. The NC model and its PCB and TICC equivalents are demonstrated to yield nonphysical overshoots in the collector-current response to sudden large signal changes in voltage excitation. The L model does not have this deficiency. When combined with the implicit-trapezoidal method the UDCS approach yields nonphysical oscillations. Substituting the Gear-2 method corrects this problem.

I develop and implement a test for the self-consistent validity of the quasi static approximation in bipolar (and MOSFET) circuit simulations to assess the relevance of NQS effects in practical digital circuits. I also investigate lateral p/n/p transistors used in a voltage regulator and in a high voltage operational amplifier. Only the L model predicts the correct amount of internal compensation required to ensure stability of the voltage regulator, and the stability of the operation amplifier. other models predict instability.

CHAPTER ONE
INTRODUCTION

Most existing models for circuit simulation, such as the Gummel-Poon model [Gum70], rely on the quasi-static approximation which in effect assumes that electrons and holes propagate with infinite velocity. This approximation misses the detail of the transient response associated with the time required for minority carriers to propagate in any direction within a quasi-neutral region. Thus, for example, in the large-signal time domain (digital circuits) delay and loading are incorrectly described by such models; in the small signal frequency domain, even excess-phase extensions of commonly used models fail to describe accurately phase margin and stability of wide-band amplifiers.

In this dissertation we focus on methods for developing and assessing new circuit models for p/n junction diodes and transistors. In developing new circuit models, I first present a phenomenological approach. This approach uses the Zimmer delay time to study the delay associated with minority carriers in quasi-neutral layers. In addition to providing better understanding of the delay phenomenon, this approach yields improved models for p/n junction diodes. I then present a systematic approach to include the effects of propagation delay. In this approach I proceed directly from the Shockley equations by a method that employs truncation in the complex-frequency domain to include the effects of the carrier propagation delay. The method

produces simple equivalent circuits that model the carrier propagation delay by an inductance (L) or a negative capacitance (NC) interacting with a resistance that models the convection current in the base. For both the small- and large-signal representations, there exists a hierarchy of model complexity. I demonstrate this systematic approach for a one dimensional quasi neutral base in low injection with an exponential doping profile, providing the framework that enables its application to [Wu89] to arbitrary doping profiles: high and moderate injection, transient or static emitter crowding and current induced base widening.

The quasi-static approximation was first used, although implicitly, in the charge-control analysis of the semiconductor devices [Roe54, Roe57, Joh58, Mid59]. This approximation was not explicitly discussed until Hamilton and coworkers [Ham63]. Recently, several one-dimensional bipolar transistor models [Fos84, Kle87, Che87] were proposed to reduce inaccuracies resulting from the quasi static approximation. I shall compare these various non quasi-static improvements against those developed in this dissertation and against the Gummel-Poon model.

The remainder of this dissertation deals with implementing the newly developed models in circuit simulators and with assessing their relevance for practical analog and digital circuits. To implement the new circuit models in SPICE for large-signal dynamic applications requires either rewriting the code or using the user-defined-control-source (UDCS) approach [Vee86, Yua87]. I employ the UDCS approach.

To address from a practical viewpoint the need for non-quasi-

static models. I develop a test for the self consistent validity of the quasi static approximation of device models that is easily implemented in circuit simulators. I demonstrate the usefulness of this test for practical digital circuits. For practical small signal applications, I investigate the need for NQS models for both a voltage regulator and a high voltage operational amplifier employing lateral p/n/p devices. High frequency microwave circuits are reviewed to study the relevance of NQS models.

I begin in Chapter Two by listing general equations and the approximations that commonly are used for bipolar semiconductor devices. I then focus on the physics of the delay phenomenon in p/n junction devices and review related earlier research studies.

In Chapter Three, I improve and exploit Elmore's delay time in determining the delay time in transient response. Since Elmore's definition is valid only for monotonic response and since the Elmore delay time is difficult to measure, I develop two new analytical methods, the area and ramp approaches, to overcome the drawbacks of Elmore's definition. The advantage of the area method over Elmore's definition is that it applies to non-monotonic responses. The ramp method provides a simple measurement technique to determine the delay time. As an application, I determine delay times associated with the convection current and the storage current of a junction diode. Taking the delay into account, I improve the quasi static model for a junction diode.

A systematic derivation of non quasi static small signal models for semiconductor junction transistors and diodes is presented in

Chapter Four. This systematic derivation proceeds directly from the solutions and introduces an inductance or a negative capacitance to account for the delay. A detailed comparison is presented by contrasting predictions of the resulting non quasi static model with exact solutions and with experiment.

In Chapter Five, I include the non quasi static propagation delay directly from the Shockley equations. I use a systematic two port approach to develop the large-and small-signal L and NC models. I discuss and compare this two port approach with previous work by Lynn et al. [Lyn67], Seltchik et al. [Sel87] and Engl and Kleostelldis [Eng69].

In Chapter Six, I compare various non quasi static improvements of the Gummel-Poon integral charge control model for both small-signal and large-signal applications. The comparison includes the L model, the NC model, the partitioned charge-based model (PCB) by Fossum and Veeraraghavan [Fos86], and the transient-integral charge-control relation (TICC) by Klose and Wieder [Klo87]. A possible fundamental limitation of the PCB, TICC and the NC models is presented.

Although the new L and NC models named above work directly in SPICE for small signal applications, their implementation in SPICE for large-signal dynamic applications either requires rewriting the code or use of the user defined control source (UDCS) approach. In Chapter Seven, I explore the UDCS approach and discuss numerical aspects of the implementation and compare different numerical integration techniques. As an application, a single transistor inverter is simulated using both the L and the NC models.

In Chapter Eight, I explore the significance of the non quasi-static (NQS) effects in circuit simulation. I propose a self-consistent validity test of the quasi-static approximation, implement it in SPICE, and apply it to a practical digital circuit.

To study the NQS effect in small-signal practical applications, I investigate (1) lateral p/n/p transistor as a pass transistor in a series voltage regulator, (2) lateral p/n/p transistor used for dc level shifting in a high voltage operational amplifier. Only the L model predicts the correct amount of internal compensation required to ensure stability of the voltage regulator and the stability of the operation amplifier, other models predict instability. I also review previous works in microwave applications to illustrate the relevance of non-quasi-static models.

Chapter Ten summarizes the contributions of this dissertation and presents recommendations for extension of the present study.

CHAPTER TWO
REVIEW OF EARLIER WORK ON TRANSIENT AND AC MODELING OF THE QUASI-NEUTRAL REGIONS

In Section 2.1, I list general equations and the approximations that commonly are used for bipolar semiconductor devices. The consequences of each approximation are discussed so are opportunities for modification. I note that Sah's transmission-line model [Sah70] demonstrates that the general (Shockley) equations can be represented by transmission-line equivalent circuits without introducing inaccuracy. Various equivalent circuit models to be developed in this study will be compared with results obtained using Sah's approach; an approach he has called circuit techniques for semiconductor analysis (CTSA).

I then study intrinsic delay in p/n junction devices and its relation to the various current components in bipolar junction devices. Phenomenological reasoning will be used to help identify the current components that contribute the delay.

2.1 List of General Equations and Approximations That Commonly Are Used for Bipolar Semiconductor Devices

General equations. The basic equations are a set of differential equations that govern the transport and recombination-generation trapping of electrons and holes in a semiconductor. They were first used by Shockley [Sho49] for getting the characteristics of p/n junction diodes and transistors. The basic equations were then

generalized by Sah [Sah71] to include optical and other generation recombination processes These equations together with auxiliary conditions describe the behavior of electrons and holes in semiconductor devices under the influence of external excitation The Shockley equations are as follows

$$\vec{J}_N = q\,\mu_n\, n\,\vec{E} + q D_n\,\nabla n \tag{2.1}$$

$$= -q\,\mu_n\, n\,\nabla v_N \tag{2.1a}$$

$$\vec{J}_P = q\,\mu_p\, p\,\vec{E} - q D_p\,\nabla p \tag{2.2}$$

$$= -q\,\mu_p\, p\,\nabla v_P \tag{2.2a}$$

$$\partial n/\partial t = \nabla\,\vec{J}_N/e - r_N + g_{(external)} \tag{2.3}$$

$$\partial p/\partial t = -\nabla\,\vec{J}_P/e \quad r_p + g_{(external)} \tag{2.4}$$

$$\nabla^2 v = -\rho/\epsilon = e(p - n + N_D^+ - N_A + \Sigma\ (trap\ charge))/\epsilon \tag{2.5}$$

$$\nabla \times \vec{E} = -\partial\vec{B}/\partial t = 0 \quad or \quad -\nabla v = \vec{E} \tag{2.6}$$

The symbols used in this dissertation are defined in the List of Symbols

Additional equations are required to describe carrier recombination, generation and trapping in semiconductor devices As an example, assume a simple one energy level Shockley Read Hall (SRH)

center. Then a seventh equation, the rate equation for the electron (or holes) trapped at the SRH center, is

$$\frac{dn_T}{dt} = [C_n(n + n_1) + C_p(p + p_1)]n_T - N_{TT}(C_n n + C_p p_1) \qquad (2.7)$$

See Sah [Sah71] for the meaning of symbols in (2.7). Sah [Sah71] has expressed these equations in their most general form to unify Impact Auger, Shockley Read Hall and other mechanisms. (He also provided an equivalent circuit model for large signals and a linear version of the model for small signals [Sah71].)

The auxiliary conditions commonly used are

$$n = \int_{E_c}^{\infty} g_c(E) \frac{1}{1 + \exp(E - E_{Fn}/kT)} dE \qquad (2.8)$$

$$= n_1 \exp[(E_{Fn} - E_1)/kT]$$

$$p = \int_{-\infty}^{E_v} g_v(E) \frac{1}{1 + \exp[(E_{Fp} - E)/kT]} dE \qquad (2.9)$$

$$= n_1 \exp[(E_1 - E_{Fp})/kT]$$

$$\nabla \cdot (\vec{J}_{CONVECTION} + \partial D/\partial t) = \nabla \cdot (\vec{J}_N + \vec{J}_P + \partial \vec{D}/\partial t) = 0 \qquad (2.10)$$

where $g_c(E)$ and $g_v(E)$ are the density of states for the conduction band and valence band respectively

$$n_i^2 = N_c N_v \exp(-E_g/kT) \qquad E_i = \frac{E_c + E_v}{2} + \frac{kT}{2} \ln(N_v/N_c)$$

Auxiliary equations (2 8) and (2 9) connect the carrier densities and the quasi-Fermi levels, which can be related to the voltage at the device terminals. Equation (2-10) states that this total current consisting of conduction and displacement components is solenoidal In a one dimensional model, this means that total current is independent of x

To assess the possibility of solution I now count unknowns and equations, in a one dimensional model for simplicity The continuity equations (2 3) and (2 4) introduce four unknowns the electron and hole densities and current densities g(external) is assumed known here If I deal with the current densities as expressed in terms of drift and diffusion components, I introduce one more unknown the electric field $\vec{E}(x) = -\nabla v(x)$ If $\vec{E}$ is known in $\vec{D} = \epsilon \vec{E}$ Eq (2 5) then becomes Poisson's equation for v(x), the fifth equation in five unknowns Since I neglect magnetic fields throughout the work, Eq (2 6) is listed for completeness

Approximations To facilitate the analysis of the bipolar junction transistor we shall make use of the following assumptions

1 Quasi-neutrality is valid in the quasi-neutral base layer Applying quasi-neutrality replaces Eq (2 5) the Poisson equation, by

$$p + \Delta p - (n + \Delta n) = N_D - N_A \quad \text{and} \quad \Delta n = \Delta p$$

for negligible charge on traps Here Δ signifies an excess density This is the critical assumption in the ambipolar approach [Van53] which provides a differential equation for $\Delta n = \Delta p$

2 Isothermal is assumed throughout the analysis

3 $\mathcal{E}_{external} = 0$

4 μ and D are not functions of frequency which requires that the excitation frequency be well below the scattering frequency [Sah73]

5 The thermal equilibrium net majority-carrier density in the base layer is essentially equal to the impurity concentration there (also an implication of quasi-neutrality) if full ionization is not assumed, additional equations describing the relationship between N_D^+ and N_D, N_A^- and N_A are required which makes analytical solution more difficult

6 No heavy doping effects is assumed For impurity densities larger than 10^{18} cm^{-3} in silicon the heavy doping effects such as energy gap narrowing majority carrier degeneracy, band tailing and impurity banding start to occur The major influences of the heavy doping effects on the basic set of equations are (a) the modification of n_1 (b) the modification of Einstein relation (c) the use of Fermi Dirac Statistics and (d) different electric fields exerted on the electron and hole For details see Marshak and coworkers [Mar81]

In addition to the preceding approximations we will occasionally simplify analyses by making the following assumptions In Chapter Five, relaxation of assumptions (7)-(10) is discussed

7 Low-injection is assumed In low injection the electric field in

(2 1) and (2 2) equals the built in electric field which is $kT/q\ d[\ln(N_{AA}/n_i)]/dx$ Substituting this electric field into (2 3) and (2 4) we obtain a linear differential equation for excess minority carrier Low-injection condition also enables the use of simpler physics to describe generation and recombination events

8 Minority current flow in the emitter region is assumed to be negligible, i e , the emitter efficiency is assumed to be unity

9 The impurity profile in the base region is uniform (or exponential)

10 One-dimensional current flow is assumed

2 2 Physics of the Delay Phenomenon in P/N Junction Devices

The electrode current of an active device has two components convection current and displacement current The convection current in a bipolar junction device is connected to the carrier injected into the principal region of the device, assumed to be the quasi-neutral base region in this study The base convection current, $j_{CONVECTION}$ in (2 10) is the sum of the electron and hole currents, each of which is a product of a particle density and velocity This finite velocity introduces delay in the time domain and phase shift in the frequency domain The displacement current component, $\partial D/\partial t$ in (2 10) of the base current affects the response only for time intervals of the order of a dielectric relaxation time Thus, practically no delay is associated with the displacement-current flow of the majority carriers

For a diode or the input part of a transistor if only majority-carrier displacement current and minority-carrier convection current were present in the principal region (quasi-neutral base) the measured

admittance of the diode would be purely conductive (Ordinary measurements would not detect influences of the majority carrier displacement current because its characteristic frequency is of the order of a reciprocal dielectric relaxation time.) The measured admittance however, is not purely conductive rather it possesses a capacitive susceptance associated with the storing of minority carriers in the volume of the base region and thus for an n-type base region in low injection, associated with the term, $(\partial p/\partial t)\delta a$, of the continuity equation (2.4) This term is proportional to $\partial v/\partial t$ for incremental changes v in the applied voltage

Note that the delay is also a central feature of such two-port devices as transistors and transmission lines One can visualize the delay effects easily in an rc transmission line in which the charging of an infinitesimal capacitor through an infinitesimal resistor at position x must precede the charging of the capacitor at position $x + \delta x$ The distributed displacement and convection currents influence the current response to the voltage at the input Similar physics governs the base region of a p/n junction device in which the storage of minority carriers at position x must occur before diffusion can accomplish such storage at position $x + \delta x$ This physics is described mathematically by the continuity equation for minority carriers (2.4)

For general time varying excitation these above equations in Section 2.1 defy solution from an analytical viewpoint Therefore device analysis rests on various schemes of approximation For example, rather than trying to deal with boundary value problems directly based on these equations one deals conventionally with

simpler, related boundary-value problems. These problems pertain to the dc steady state in which the time varying terms in the above equations are set equal to zero. By various strategies of approximation, one then solves these equations in part, and obtains useful relations. For example, for dc conditions one finds expression for the charge (q) that resides in different regions of the device. These charges change with time if a net influx of mobile carriers occur from the various terminals of the device. Thus, as a first step in the quasi-static method of analysis, one finds the functional dependence of each of these charges on the dc voltages at the terminals, and finds similar expressions for the current. Then, as the central step, one makes the quasi-static approximation that these functional dependencies for the charges and the currents apply not only in the dc steady state but also during transient conditions. That is, if

$$q = q(V_1, V_2, V_3, V_4), \text{ etc} \tag{2.11}$$

for the dc steady state, then one assumes that

$$q(t) = q[v_1(t), v_2(t), v_3(t), v_4(t)] \tag{2.12}$$

during transients.

Often in the modeling of p/n junction devices, one typically uses the quasi-static approximation to enable a simple description of the storage current in quasi neutral regions. The storage current is $dQ/dt = dQ/dv\ dv/dt$, where the excess carriers stored in a quasi neutral

region $Q = q\int An_p dx$ are assumed for transients to have the same functional dependence on voltage v that they have in the dc steady state. This approximation implicitly assumes, for example, that the excess electrons injected in a quasi-neutral p-type region move with infinite velocity to produce during transients, a steady state distribution in position x at each time t. Thus it neglects the electron propagation delay associated with its convection current. In the quasi-static approximation, delay is neglected. This will receive close attention in this study.

2.3 Related Earlier Research

Many p/n junction device models are based on the quasi-static approximation that the minority carrier space dependence in the quasi-neutral base retains its steady state form even during transients. This approximation neglects the distributed nature of the devices and the minority-carrier propagation delay associated with the convection current.

2.3.1 Bipolar Junction Diodes

The quasi-static diode capacitance dQ/dV calculated by the quasi-static approximation in the quasi-neutral base does not agree with the exact capacitance calculated from the susceptance, $Im\{Y(j\omega)\}/\omega$, where Y is the small-signal admittance. This disagreement persists even in the limit as $\omega \to 0$ [Che71, Van75, Bul68].

Since the charge control capacitance $\partial Q/\partial v$ takes no account of the convection current, Cherry [Che71] noted that the driving-point (or

measured) capacitance. C_{ln} is less than the charge control capacitance, C_{qs}, because the convection current contributes a negative susceptive component to the admittance. Van der Ziel [Van76] calculated both the C_{ln} and C_{qs} for a long-base p/n diode, determining thereby $C_{ln} = C_{qs}/2$. He then associated the convection current with a phase shift to explain the difference between the two capacitances.

Schlueter [Sch65] wrote a computer program based on the transmission-line representation of the internal processes in the n-type neutral region of a long-base p/n diode. This program calculated the actual capacitance and the quasi-static capacitance, which ignores (in the low frequency range) the presence of the capacitor when calculating the voltage along the line. He showed that the quasi-static approach introduces a systematic error by neglecting the distributed nature.

Lindmayer and Wrigley [Lin64] explained the difference between the C_{ln} and C_{qs} of a p/n diode by studying the transient currents in the base region. Observing the turn-off transient of a diode in which the forward voltage drops to zero, Lindmayer and Wrigley demonstrated that the stored holes diffuse in both directions as time passes (Fig. 2-15 in Lindmayer [Lin64]). As a result, the opposing currents will move less net hole charge through the external circuit. Mathematical calculation determined that only 1/2 and 2/3 of the stored hole charge is reclaimable through the junction for a long base diode and a short-base diode, respectively. Since the driving point capacitance is half and two-thirds of the quasi-static capacitance for long-base and short-

base diodes, respectively. Lindmayer and Wrigley connected the idea of reclamation of charge to explain the difference between the two capacitances.

2 3 2 Bipolar Junction Transistors

Giacoletto [Gia54] calculated susceptive components for both trans-admittance and self-admittance parameters of a transistor. He employed a heuristic consideration of the carrier transit time to approximate the diffusion portion of the admittance. Giacoletto did not explain how he obtained the value of the carrier transit time, $X_B^2/6D_p$. He may have obtained that value by knowing what phase angle was required to produce the correct $Im[Y(j\omega)/\omega]$. Chen and coworkers [Che87] had the correct input description of the quasi-neutral base. They obtained a model containing an inductor which approximates the influence of propagation delay by filtering out high frequencies.

To improve accuracy above that obtainable from the quasi-static approximation is bipolar transistor modeling, Feesum and Veeraraghavan [Fea86] and Kiese and Wieder [Kie87] developed the partitioned-charge based (PCB) model and the transient integral-charge-control (TICC) model, respectively. Detailed discussions of these two models are given in Chapter Six.

Seltchik et al. [Sel87] developed a new small-signal model, which improves the transient characterization of bipolar transistors by combining the ideas from the previous high and low frequency circuit representations. A discussion of Seltchik's model is given in Section

5 6 1 In Chapter Five I also review previous work by Lynn et al. [Lyn67] and by Engl and Kioustelidis [Eng69]

CHAPTER THREE
IMPROVED CALCULATIONS OF DELAY AND THEIR APPLICATION TO BIPOLAR JUNCTION DIODES

The delay time is a key feature of the transient response of two-port linear systems. The conventional definition of 50 per cent delay time is commonly used. However, exact calculations of this delay time are complicated. A convenient analytic method for computing the delay time which applies for a monotonic (non overshooting) step response is given by Elmore [Elm48]. The definition used in effect by Asher to analyze switching transistor circuits [Ash64].

Elmore's definition of delay time τ_D is as follows [Elm48; Ash64]

$$\tau_D = \frac{\int_0^\infty t\, e'(t)\,dt}{\int_0^\infty e'(t)\,dt} = \lim_{s\to 0} \frac{-\frac{d}{ds}\,[E'(s)]}{E'(s)} \qquad (3\text{-}1)$$

where $e'(t) = d/dt[e(t)]$ is the impulse response (either current or voltage) of the network and $E'(s)$ is its Laplace transform.

The physical interpretation of Elmore's delay time can be illustrated by considering minority-carrier propagation in a quasineutral layer of a p/n junction diode. For example, one may think of Fig. 3-1(a) as representing the flux of minority carriers that have

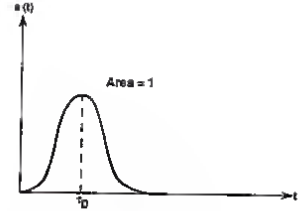


Fig 3 1 Typical transient response curves
(a) Normalized impulse response, (b) Normalized step response

propagated across the base in time t in response to a step-function excitation of minority-carrier density at the left side of the base ($x = 0$). In this example, $e(t)$ may be regarded as the probability that carriers inserted at $x = 0$ reach the right side of the base ($x = X_B$) in time t. This probability interpretation follows because carrier diffusion and drift are random (stochastic) processes. Thus $e(t)$ may be regarded as a cumulative probability function and $e'(t)$ as a probability density for the random variable t. Elmore did not consider this interpretation, but we shall make use of this view of the Elmore definition in much of the rest of this work.

The Elmore's definition of delay time has several advantages. First, it follows directly from the transfer function of the system. Second, it has the property that the overall delay is the sum of the delays of the individual stages. But the Elmore delay time applies only to transfer functions that yield monotonic step responses in the time domain and no procedure exists for measuring it directly.

We develop here two new methods for calculating the Elmore delay time and illustrate them by application to p/n junction devices. These methods enable application to non-monotonic step responses and are convenient for measurement.

3.1 Area Approach

Consider a step function excitation. The Elmore delay time is the ratio of the shaded areas in Fig. 3.2 to the final value of the response [Che58].

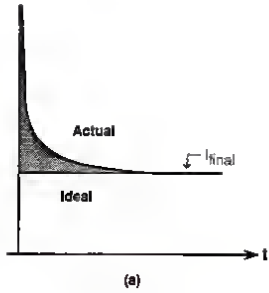


Fig 3 2 Comparison between ideal and actual step responses of a p/n junction diode (Ideal means absence of delay)
(a) Overshoot step response (b) Monotonic step response

$$\tau_{D(\text{area})} = \frac{\int_0^\infty \{(\text{ideal} - \text{actual}) \text{ response}\ dt}{I_{final} \text{ or } V_{final}} \qquad (3\ 2)$$

as is proved below. The ideal response in the denominator of (3 2) is defined as the output response that follows the input excitation without delay.

The overshoot in the current response (Fig. 3 2), which occurs for example in a p/n junction diode (one port) excited by a step-function voltage, makes the Elmore definition inapplicable, but it is accommodated by the definition given in (3 2). Applying (3 2) to the actual response with overshoot (Fig. 3 2a) results in a negative-valued delay time, which correctly corresponds to a lead in phase angle. Delay corresponds to lag in the frequency domain. Lead corresponds in a mathematical sense to negative delay. But no implication should be construed that the system is nonsensual.

Equation (3 2) is mathematically equivalent to (3 1), which we now prove. The Elmore definition of delay time for step response $e(t)$ is

$$\tau_{D(\text{Elmore})} = \frac{\int_0^\infty t\, e'(t) dt}{\int_0^\infty e'(t)\, dt} = \frac{\int_0^\infty t\, [d(e(t))/dt]\, dt}{\int_0^\infty [d(e(t))/dt]\, dt} \qquad (3\ 3)$$

Partial integration yields

$$\tau_{D(\text{Elmore})} = \frac{\int_0^{\infty} e(t)\,dt}{e(t=\infty) \;/\; e(t=0)} \qquad (3.4)$$

The infinity appearing in the improper integral above signifies the usual limit operation. Thus we may substitute a large enough value T_0 for ∞ that $e(T_0)$ and $e'(T_0)$ are arbitrarily close to E_0 and 0 respectively. Thus

$$\tau_{D(\text{Elmore})} = \frac{\int_0^{T_0} e(t)\,dt}{E_0} \qquad (3.5)$$

Since E_0 can be defined as either V_{final} or I_{final} in Fig. 3.2, Eq. (3.5) is identical to (3.2).

3.2 Ramp Approach

The ramp approach generalizes Elmore's definition. The generalized definition enables treatment of non-monotonic step responses. But measurement of $\tau_{D(\text{Elmore})}$ requires the integration shown in Fig. 3.2, which suggests that we explore a ramp rather than a step response. Typical ramp responses (Fig. 3.3a and 3.3b) enable direct measurement of delay.

In the ramp approach, the delay time equals the difference between the starting time and the time extrapolated from the ramp response back to the time axis (Fig. 3.3a and 3.3b). It is also equivalent to measuring the time difference between the ideal and actual response

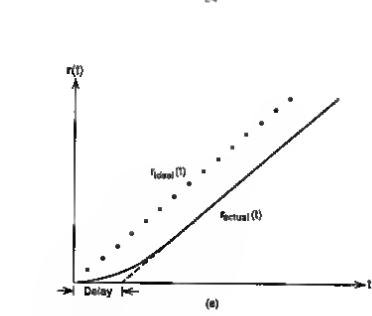


Fig 3 3 Some typical ramp response curves
(a) Typical ramp response with delay and an ideal response
(b) Typical ramp response with lead and an ideal response.
(c) The alternate method to measure the delay time

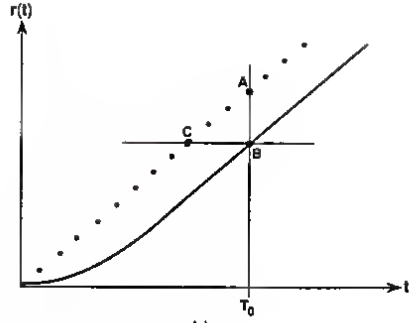


Fig. 3.3 (continued

after some time T_0 which is large enough to allow the responses to follow the excitation (Fig 3 3c) BC which is defined in Fig 3 3(c) is independent of the slope of the ramp input because

$$BC = AB/slope = AB/E_0 = [E_0 T_0 - \int_0^{T_0} e(t')dt']/E_0 \qquad (3\ 6)$$

which is identical to (3 5) and hence to (3 2)

3 3 Applications

Here we demonstrate applications of the analytical determination of delay time We shall first calculate the delay associated with the convection current using the Elmore's definition Since the quasi static based model neglects the minority-carrier propagation delay associated with the convection current (see Chapter Two), we modify quasi-static small signal models for the diode by taking this delay into account This modification provides two improved models one has an inductor L in series with the resistor R that models the convection current the other has a negative capacitor in parallel with the resistor R The modified models describe the frequency and transient response more accurately

The next higher-order modification is obtained by inserting the delay associated with the minority-carrier storage charge in the base using the Elmore's definition This is the topic of Section 3 3 2

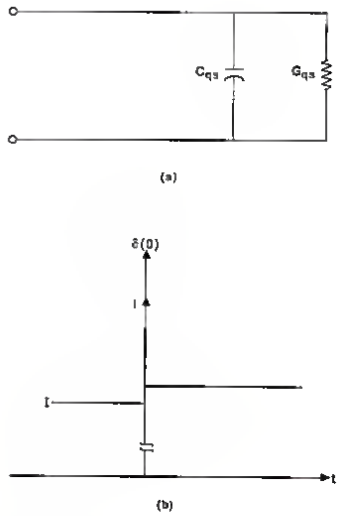


Fig 3 4 A simple quasi-static model for a p/n junction diode and its response to a small step voltage excitation.
(a) The quasi-static model (b) The response

3 3 1 Delay in the Convection Current

The quasi-static small signal model shown in Fig 3 4(a) yields the response to small step voltage excitation that is illustrated in Fig 3 4(b) This response consists of an impulse followed by a time-independent current The impulse is the storage current which flows through C_{qe} in the equivalent circuit, and the time independent current, which flows through G_{qs}, is the convection current The quasi-static model is inaccurate because it neglects the delay, $\tau_{D(c)}$ associated with the convection current

To improve this model, we now insert this delay in the time domain Let the quasi-static current through G_{qe} be $i_{qe}(t)$ Then the current

$$i(t) = i_{qe}(t)\, u(t - \tau_{D(c)}) = G_{qs}\, v(t)\, u(t - \tau_{D(c)}) \qquad (3\ 7)$$

is the convection current modified to contain the propagation delay Taking the Laplace transform of this equation yields

$$I(s) = G_{qs} \exp(-s\tau_{D(c)})\, V(s) \qquad (3\ 8)$$

which we represent symbolically by Fig 3 5(a) The exponential is inconvenient for computer-aided analysis so we approximate it as

$$G_{qs} \exp(-j\omega\tau_{D(c)}) = G_{qs}/\exp(j\omega\tau_{D(c)}) \approx G_{qs}/(1 + j\omega\tau_{D(c)})$$

$$= 1/(1/G_{qs} + j\omega\tau_{D(c)}/G_{qs}) \qquad (3\ 9)$$

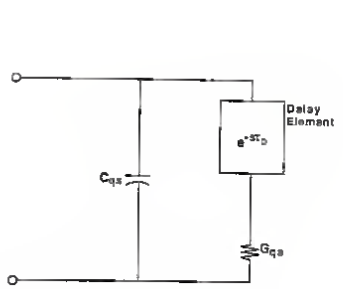
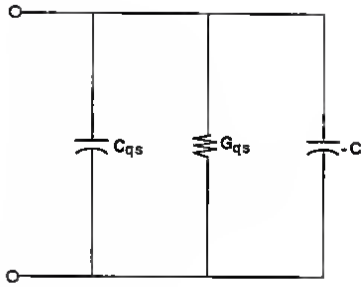


Fig 3 5 Modified p/n diode models taking into account the phase shift in the convection current
(a) A symbolic model, (b) The DL-1 model, (c) The DNC-1 model



(c)

Fig 3.5 -continued

This yields the equivalent circuit of Fig. 3.5(b), in which $R_{qs} = 1/G_{qe}$ and $L/R_{qe} = \tau_{D(c)}$. An alternate expansion yields the equivalent circuit of Fig. 3.5(c), in which $G_{qe}/C_{qe} = \tau_{D(c)}$.

$$G_{qs} \exp(-j\omega\tau_{D(c)}) \approx G_{qe}(1 - j\omega\tau_{D(c)}) = G_{qe} + j\omega(-C) \qquad (3.10)$$

We name the modified model in Fig. 3.5(b) the DL-1 model and the modified model in Fig. 3.5(c) the DNC-1 model.

To complete the modified models (Fig. 3.5b, 3.5c) we must determine $\tau_{D(c)}$. As an illustration, we focus on the p-type quasi-neutral base region of a p/n junction diode, assuming low injection, negligible recombination, and a position-independent doping density. From the quasi-static model of Fig. 3.4(a), using the area approach (Eq. 3.2) we obtain that the delay time is the negative of the quasi-static transit time $X_E^2/2D_n$, which is the quasi-neutral base thickness squared divided by twice the electron diffusivity. From the standard exact solution for the step response of a diode [Van76], using either (3.1) or (3.2), we determine that the delay time is two-thirds of that of the quasi-static model. The exact solution includes the effect of propagation delay associated with convection current and the quasi-static model does not. Therefore, the difference between these two delay times yields the delay time associated with the convection current:

$$\tau_{D(c)} = (1/3)\, X_E^2/2D_n \qquad (3.11)$$

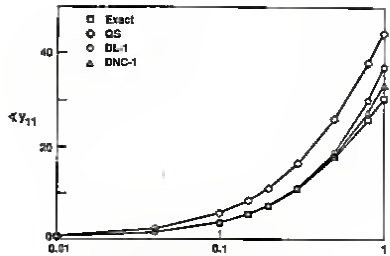
Equation (3 11) can also be obtained by an alternate approach in which a second part is introduced artificially [Che86]

The preceding discussion has assumed that the quasi-static transit time is much less than the recombination lifetime τ_n For a semi infinite or long-base diode in which all injected electrons recombine with holes similar calculations produce

$$\tau_{D(c)} = (1/2)\,\tau_n \qquad \text{if} \quad X_B^2/2D_n \gg \tau_n \tag{3 12}$$

For the short-base diode the delay associated with the convection current arises from the time required for carriers to cross the quasi-neutral base For a long base diode the recombination current is the convection current component Its delay arises from the coexistence during transients of the storage current and the recombination current

In Figure 3 6, we compare the input admittance y_{11} of a short base diode predicted by the DL-1 model (Fig 3 5b) the DNC 1 model (Fig 3 5c) and the quasi static models (Fig 3 4a) The standard for comparison is the exact y_{11} for an uniformly doped base The phase of the small-signal $y_{11}(\omega)$ is simulated well by both the DL-1 and DNC 1 models whereas the simulations of the quasi-static model are inaccurate even at very low frequencies Further, for arbitrary base thickness the modified models predicts the correct input capacitance The DL-1 and DNC 1 models yield less improvement for the magnitude of y_{11}



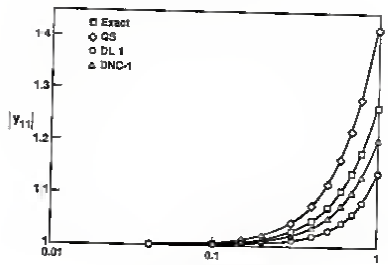


Fig. 3 6 Input admittance of a short base diode predicted by the quasi-static model the DL 1 model and the DNC 1 model The standard for comparison is the exact solution
(a) Phase (b) Magnitude

The DL-1 and DNG-1 models can also be derived by approximating thenetwork representation of the minority-carrier continuity equation (a two-port network) This is the subject of Chapter Five

3 3 2 Delay in the Storage Current

In the last section we derived the DL-1 and DNG-1 models by inserting the delay time associated with the convection current $\tau_{D(c)}$ to the quasi-static small signal model Here we further improve the DL-1 and DNG-1 models by inserting the delay associated with the storage current $\tau_{D(s)}$ The delay associated with the storage current arises from the finite time required for minority carriers to build up in the quasi-neutral base For example if a diode is subjected to a step increase in the input voltage the storage charge delay results because the storage charge does not instantaneously reach its final value

We now insert $\tau_{D(s)}$ in the time domain From the discussion in the last section, the storage current is the current flowing through C_{qs} Then the current

$$i(t) = C_{qs} v(t) u(t - \tau_{D(s)}) , \tag{3 13}$$

is the storage current modified to contain the finite time delay Taking the Laplace transform of (3 13) yields

$$I(s) = C_{qs} \exp(-s \tau_{D(s)}) V(s), \tag{3 14}$$

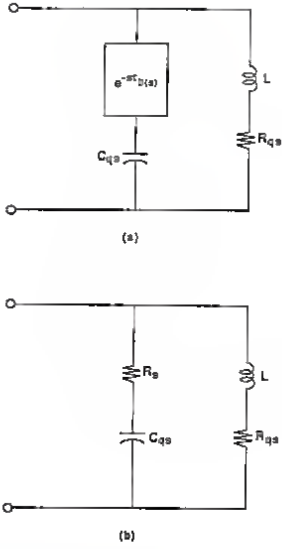
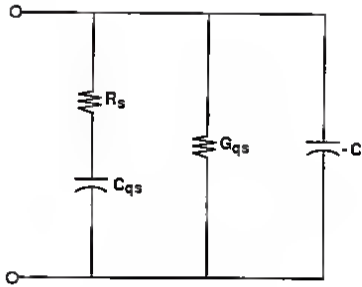


Fig 3 7 Modified p/n diode models taking into account the phase shift in the storage current
(a) A symbolic model, (b) The DL-2 model (c) The DNC-2 model



(c)

Fig 3 7--continued

which we represent symbolically by Fig 3 7(a) The exponential is inconvenient for computer-aided analysis so we approximate it as

$$C_{qs} \exp(-j\omega\tau_{D(s)}) = C_{qs}/\exp(j\omega\tau_{D(s)}) \approx C_{qs}/(1 + j\omega\tau_{D(s)})$$

$$= C_{qs}/(1 + j\omega C_{qs} R_s) \qquad (3.15)$$

This yields the equivalent circuit of Fig 3 7(b), in which $R_s = \tau_{D(s)}/C_{qs}$

To complete the modified model (Fig 3 7b) we must determine $\tau_{D(s)}$ We again focus on the p type quasi neutral base region of a p/n junction diode, assuming low injection negligible recombination and a position-independent doping density To determine the delay time associated with the storage charge we solve the minority-carrier continuity to obtain the excess minority carrier density in x and s Laplace transformation of the minority carrier continuity equation yields a second-order differential equation in position x, in which the complex frequency variable s is treated as a parameter The solution of this ordinary differential equation for the transformed excess electron density involves two arbitrary constants We remove these constants by applying the boundary conditions at the two edges of the quasi-neutral base layer Thus, we obtain the excess electron density

$$n(x, s) = n(0,s) \sinh\left(\frac{X_B - x}{L_n^*}\right)/\sinh\left(\frac{X_B}{L_n^*}\right) \qquad (3.16)$$

where n(0, s) is the initial condition at x = 0 $L_n^* = L_n/(1 + s\,\tau_n)^{1/2}$ for a long-base diode, and $L_n^* = (D_n/s)^{1/2}$ for a short base diode Integrating (3-16) from x = 0 to x = X_g yields the expression for the storage charge

$$q_{ST}(s) = \frac{n(0\ s)\ L_n^*}{\sinh(\frac{X_g}{L_n^*})}\ [\cosh(\frac{X_g}{L_n^*}) \quad 1] \tag{3-17}$$

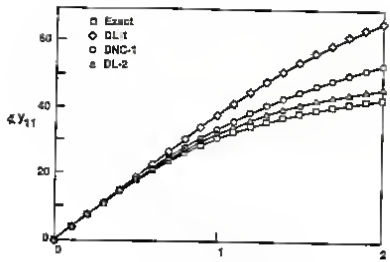
Applying (3 1) to (3 17), we determine the delay time associated with q_{ST} as follows

$$\tau_{D(s)} = \tau_{tr}/6 \quad \text{for a short-base diode} \tag{3 18}$$

and

$$\tau_{D(s)} = \tau_n/2 \quad \text{for a long-base diode} \tag{3 19}$$

We apply this modification to the DL 1 and DNC-1 models and designate the modified DL 1 model in Fig 3 7(b) as the DL-2 model and the modified DNC 1 model in Fig 3 7(c) as the DNC-2 model In Figure 3 8 we compare the input admittance, y_{11}, of a short-base diode predicted by the DL 2 model with the DNC-1 and the DL 1 models The standard for comparison is the exact y_{11} for an uniformly doped base Compared with the DL-1 and DNC-1 models the DL 2 model gives much greater accuracy in simulating the phase and the magnitude of the



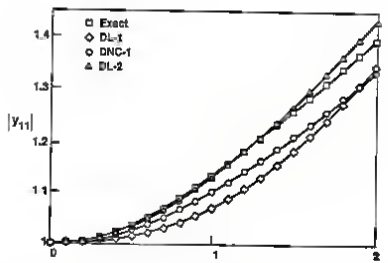


Fig 3 8 Input admittance of a short base diode predicted by the DL-1 model, the DNC 1 model and the DL-2 model. The standard for comparison is the exact solution (a) phase (b) magnitude

small signal $y_{11}(\omega)$ In Figure 3 9 we see that the DL-2 model and the DNC 2 model have similar accuracy in predicting y_{11}

Similarly to the DL-1 and DNC-1 models, the DL-2 and DNC-2 model can also be derived by approximating the network representation of the minority-carrier continuity equation (See Chapter Five)

3 4 Generalization to Large-Signal Excitations

For large signal excitation, the relation between current and voltage in the quasi-neutral base layer is nonlinear, whereas the relation between the minority-carrier density and the minority carrier current is linear for low and high injection This linear relationship between the minority-carrier density and the minority-carrier current enables the application of the Elmore delay time, which is valid only for a linear system, to large signal excitation The conversion between voltage and minority-carrier density is straightforward

Previous workers have utilized the linear relation existing between the minority carrier density and the minority carrier current in the quasi-neutral base layer For example Linvill [Lin56] and Linvill and Gibbons [Lin61] developed the Linvill model which exploited this linearity by use of elements relate minority-carrier current and density, thus avoiding the nonlinearities between current and voltage

3 5 Summary and Discussion

In this chapter we reviewed the Elmore's definition of delay time and described modifications The Elmore delay time is restricted to

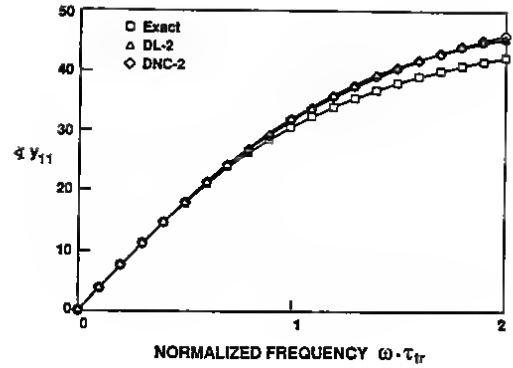


Fig 3 9 The phase of the input admittance of a short-base diode predicted by the DL 2 model and the DNC 2 model The standard for comparison is the exact solution The magnitude of y_{11} predicted by the DNC-2 model is similar to that of the DL-2 model (Fig 3 8b) and is not shown

monotonic response. The area approach enables treatment of non monotonic step response. The measurement of $\tau_{D(Elmore)}$ requires differentiation (Fig. 3.1b) or integration (Fig. 3.2). The ramp approach enables direct measurement. As an application, the small-signal equivalent circuit for a p/n junction diode based on the quasi-static approximation was modified to account for minority carrier propagation delay. The application of the Elmore's delay time to large signal excitation is discussed.

For frequencies lower than the reciprocal of the delay time, the improved models provide (a) the correct driving-point admittance, (b) the correct phase shift between the current and the voltage, and (c) better accuracy in computer-aided analysis and design [Cho86]. The modified model can also be derived by approximating the network representation of the minority-carrier continuity equation (a two-port network), which is the subject of Chapter Five. In the next chapter the modified small-signal equivalent circuit for diodes will be extrapolated to that for bipolar junction transistors. The resulting modified transistor model correctly accounts for carrier propagation delay, which is a refinement not included in the conventional hybrid-π model.

CHAPTER FOUR
SYSTEMATIC NON-QUASI-STATIC SMALL-SIGNAL MODELING FOR SEMICONDUCTOR JUNCTION TRANSISTORS AND DIODES

For small-signal and hence linear characterizations of diodes and transistors exact solutions are available [She49]. These solutions, however, become intractable for many computer-aided analysis programs because their network representation involves frequency-dependent controlled sources. For this reason, and to enhance physical understanding, the conventional small-signal diode models [Van76] and hybrid-π transistor model [Gie54] rely on the quasi-static approximation.

These conventional small signal equivalent circuits for transistors and diodes exclude effects of carrier propagation delay. In Chapter Three, I modified the quasi-static model of a diode by imbedding a delay element, which results in an inductance in the DL 1 model and a negative capacitance in the DNC 1 model.

In the alternate method for deriving the modified equivalent circuit models to be presented, we shall see that the delay element instead of being put in, in the sense that is described in Chapter Three, enters systematically from approximating network representations of the minority-carrier continuity equation. We first demonstrate this. Then I exploit the indefinite admittance matrix [She52] to transform systematically from common-base to common-emitter

configurations and finally I assess the accuracy of the non quasi static models by comparing their predictions against exact solutions and against experiment

4 1 Systematic Inclusion of Delay for Transistors and Diodes

To begin the derivation, for simplicity I assume a uniformly doped n-type quasi-neutral base region in low injection Further I assume unity injection efficiency Following the two-port approach of Lindholm and Hamilton [Lin64], I take the Laplace transform with respect to time of the minority-carrier continuity equation This yields an ordinary differential equation, the solution of which is reformulated to express the transforms of the weighted incremental voltage and current at the emitter edge (input port) of the quasi-neutral base in terms of their counterparts at the collector edge (output port) The following two-port characterization results [Che87]

$$A_{11}(s) = A_{22}(s) = \frac{A\, q\, D_p}{L_p^*} \coth\left(\frac{X_B}{L_p^*}\right) \tag{4 1}$$

and

$$A_{12}(s) = A_{21}(s) = -\frac{A\, q\, D_p}{L_p^*} \operatorname{csech}\left(\frac{X_B}{L_p^*}\right) \tag{4 2}$$

where $L_p^* = L_p/(1 + s\,\tau_p)^{1/2}$ if base recombination is negligible, $L_p^* = \{D_p/s\}^{1/2}$

Fig 4 1(a) shows a π network representation of this two port characterization Using this and truncating the parameters as follows

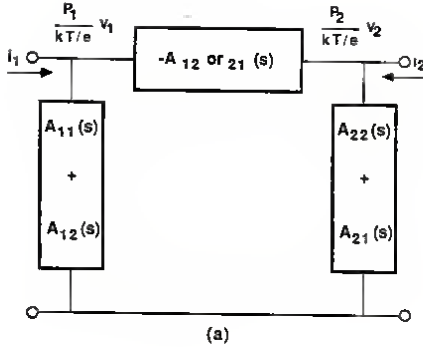
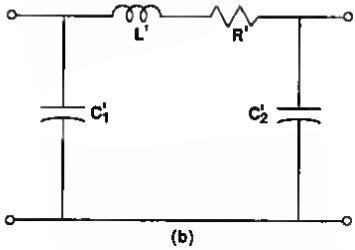


Fig 4 1 Small-signal equivalent circuit for a quasi-neutral base of transistors and diodes.
(a) The two-port network where $P_{1,2}$ are the minority carrier densities at $x = 0$ and $x = X_B$, respectively (b) After the truncation (4 4). (c) A simplified short-base model (the DL 1 model) with an ohmic contact, where $L = L'(kT/q)/F_1$ $R = R'(kT/q)/F_1$ and $C_1 = C_1' F_1/(kT/q)$ In this model, the modal variable is the small signal voltage.



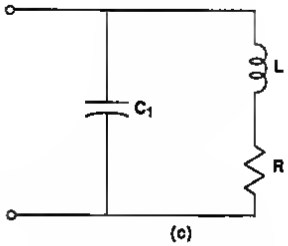


Fig 4 1 - continued

$$A_{11}(s) + A_{12}(s) = s\,A\,q\,X_B/2$$

$$-A_{21}(s) = -A_{12}(s) = \frac{A\,q\,D_p}{L_p^* \sinh(X_B/L_p^*)} = 1/\left(\frac{X_B}{A\,q\,D_p} + s\,\frac{X_B^3}{6A\,q\,D_p^2}\right),$$

yields the circuit representation of Fig 4.1(b), in which the parameter values of network-like elements are given by

$$C_1' = C_2' = A\,q\,X_B/2, \quad R' = X_B/A\,q\,D_p \text{ and } L' = X_B^3/(6A\,q\,D_p^2)$$

The preceding derivation can be easily generalized by relaxing the assumption that the base doping concentration is position independent (See Chapter Five). The appearance of the inductor L' in series with R' arises symematically from the physics.

For this network representation of the convection current, the time constant for a short-base diode is $L'/R' = X_B^2/6D_p = \tau_D$, which is the delay calculated from y_{21} by the method of Ashar [Ash64]. This indicates that the series combination represents the effect of minority-carrier propagation delay. A similar series circuit represents the convection current in the Sah transmission-line model when the excitation frequency is comparable to the scattering frequency [Sah73]. This was shown by a moment method solution of the Boltzmann transport equation.

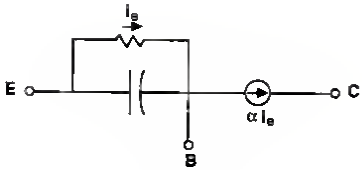
If I assume an ohmic contact at the second port, the capacitor C_2' is shorted. Putting the weighted potential, $F_{DC}(0)/(kT/q)$, into C_1', L' and R', I obtain network elements C_{qs}, L and R_{qs} and the small-signal model for a diode, shown in Fig. 4.1(c). Simple calculations

yields

$$C_1 = \tau_{tr}\, J/(kT/q) \quad R = kT/(q\,J), \text{ and } L = \tau_D\,(kT/q)/J$$

where $J = q\,D_p\,p_{DC}(0)/X_B$ is the dc component. Comparing Fig. 4.1(c) to the DL-1 model (Fig. 3.5b) I notice that not only is Fig. 4.1(c) equivalent to the DL-1 model but the value of the capacitors, resistors, and inductor in both equivalent circuit models are identical ($C_{qe} = C_1$, $R_{qe} = R$). Hence the foregoing procedure provides an alternate derivation of the DL-1 model. Since the DL-1 model gives the correct admittance at low frequencies, this network representation illustrates that the measured small-signal capacitance differs from the quasi-static capacitance because it includes minority carrier propagation delay. Thus, it enables a simple interpretation of the exact solution [Sho49] for the input admittance.

For a transistor, the network representation provides additional benefits. The simplest common-base model (Fig. 4.2a) for forward-active operation is a quasi-static T model. I see that the application of an ideal step voltage to the model of Fig. 4.2(a) yields an output (collector) current with no delay. Similarly, the collector-current response to an ideal sinusoidal voltage source at the input is independent of frequency. I now develop systematically a model that includes the effects of propagation delay in both the input and transfer characteristics. This model (Fig. 4.2b) differs from the model of Fig. 4.2(a) only by the presence of the inductor in series



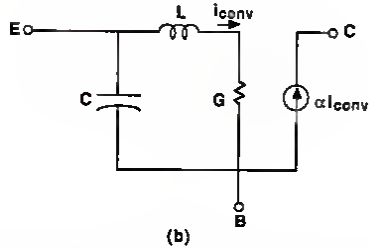


Fig 4 2 Small signal equivalent circuits for transistor in common-base mode
(a) Simplest equivalent circuit (quasi-static) (b) Modified model (c) Modified model in common-emitter mode

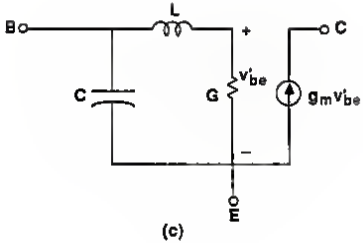


Fig 4 2--continued

with the resistance that carries the convection current in the input circuit

For a systematic derivation, I introduce the h parameters of two-port network theory [Lin61]. h_{11} is the input impedance of a short-base diode (Fig. 4-1c) and of a transistor in forward active operation. The transfer parameter h_{21} is α, the short-circuit current gain. I assume that α is independent of V_{bc} so that $h_{12} \approx 0$ [Lin61]. If I also assume no base-width modulation, $h_{22} \approx 0$. Thus I obtain the small-signal equivalent circuit for the transistor shown in Fig. 4-2b.

If the emitter rather than the base is ground, I obtain the standard hybrid-π model modified by the presence of an inductor in the branch representing convection current (Fig. 4-2c). For the common-emitter mode, i in $di/dw = y_{11}$ is the base current rather than the emitter current. Thus the ratio of $1/R_{(common\ base)}$ to $1/R_{(common\ emitter)}$ is $[di_{(base)}/di_{(emitter)}]$, which is $1/(\beta + 1)$. Because $L = r_d R_{qs}$, $L_{(common\ emitter)}/L_{(common\ base)} = (\beta + 1)$. The capacitance, which is dQ/dv_{BE}, is independent of the choice of the ground terminal. In Fig. 4-2c, I set $i = g_m v' \approx y_{21}(s) v_i$, consistently with Eq. (4-4).

As derived in Chapter Three, inclusion of minority-carrier propagation delay also can be represented by a negative capacitor in parallel with R_{qs}. This negative capacitance model (the DNC-1 model) can be systematically derived by truncating $A_{21}(s)$ in an alternate way

$$A_{21}(s) = A_{12}(s) = A \cdot D_p/[L_p^* \sinh(X_B/L_p^*)]$$

$$- A\,q\,D_p\,\mathrm{cosech}(X_B/L_p^*)/L_p^* \approx \frac{A\,q\,D_p}{X_B}\left(1 + s\,\frac{X_B^2}{5}\right)$$

Equation (4.5) yields the circuit representation of the Fig. 4.3(a), in which the parameter values of network-like elements are given by:

$$C_1' = C_2' = A\,q\,X_B/2,\quad R' = X_B/A\,q\,D_p \text{ and } -C' = -X_B^2/(6R')$$

I can then derive the DNC-1 model as shown in deriving the DL-1 model above.

4.2 Conversion among Different Configurations

I now present a systematic approach, using the indefinite admittance matrix method, to convert among different choices of reference terminals. The associated two-port matrix of the modified small-signal equivalent circuit for an intrinsic transistor in the common-base configuration (Fig. 4.2b), is

$$\begin{bmatrix} i_e \\ i_c \end{bmatrix} = \begin{bmatrix} Y_{ee} & Y_{ec} \\ Y_{ce} & Y_{cc} \end{bmatrix} \begin{bmatrix} v_e \\ v_c \end{bmatrix} \qquad (4.6)$$

where $Y_{ee} = s\cdot C + 1/(R + s\cdot L)$, $Y_{ec} = Y_{cc} = 0$ and $Y_{ce} = -s/(R + s\cdot L)$. Knowing that the sum of the coefficients in each row and in each column of the associated indefinite admittance matrix [She52] is zero, I derive for this matrix:

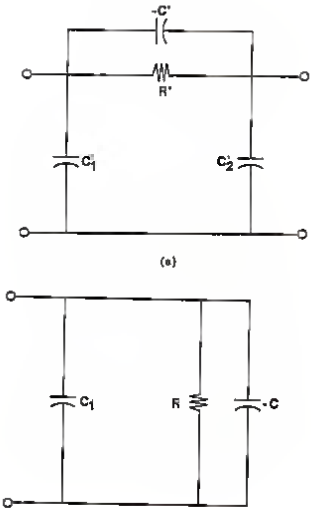


Fig 4.3 The two port small-signal equivalent circuit for a quasi-neutral base of transistors and diodes (Fig 4.1a).
(a) After the truncation (4.5) (b) A simplified short-base model (the DNC 1 model) with an ohmic contact where $C = C' F_1/(kT/q)$, $R = R'(kT/q)/F_1$ and $C_1 = C_1' F_1/(kT/q)$.

$$\begin{bmatrix} i_e \\ i_c \\ i_b \end{bmatrix} = \begin{bmatrix} Y_{ee} & Y_{ec} & Y_{eb} \\ Y_{ce} & Y_{cc} & Y_{cb} \\ Y_{be} & Y_{bc} & Y_{bb} \end{bmatrix} \begin{bmatrix} v_e \\ v_c \\ v_b \end{bmatrix} \qquad (4\ 7)$$

Any particular configuration of the transistor is obtained by reducing the 3 by 3 matrix of Eq. (4.7) to a 2 by 2 matrix by crossing out both the column and row corresponding to the common terminal in the two-port configuration. For example, for the common emitter configuration the column and row associated with emitter (that is, the first column and first row in Eq. (4.7)) are deleted leading to

$$\begin{bmatrix} i_c \\ i_b \end{bmatrix} = \begin{bmatrix} Y_{cc} & Y_{cb} \\ Y_{bc} & Y_{bb} \end{bmatrix} \begin{bmatrix} v_c \\ v_b \end{bmatrix} \qquad (4\ 8)$$

where

$$Y_{bc} = -(Y_{ec} + Y_{cc}) = 0,$$

$$Y_{bb} = -(Y_{eb} + Y_{cb}) = Y_{ee} + Y_{ce} + Y_{ec} + Y_{cc}$$
$$= s\,C + 1/(R + s\,L) - \alpha/(R + s\,L) = s\,C + 1/(R^{CE} + s\,L^{CE})$$

and

$$Ycb = -(Yce + Ycc) = \alpha/(R + s\,L) = \beta/(R^{CE} + s\,L^{CE}) = g_m\,v'$$

in which R^{CE} and L^{CE} equal $(1 + \beta)R$ and $(1 + \beta)L$, respectively.

The resulting equivalent circuit is shown in Fig 4 2(c) The common-collector configuration is similarly derived

Inclusion of extrinsic resistors is straightforward

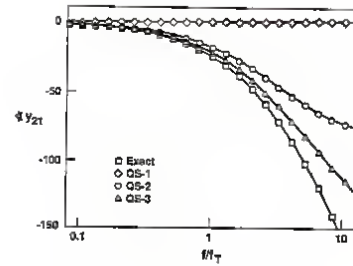
4 3 Small Signal Assessment

4 3 1 Comparison with Exact Solution

I now assess accuracy in frequency response of the two-port parameters of the models developed up to now The standard for comparison is the exact solution obtained from the partial differential equations The assessment includes higher-order models (multiple section models) derived by dividing the base region into number of subregions Applying the systematic derivation (Section 4 1) to each subregion, I construct similar models for each subregion of the base, connecting them all in tandem to form overall multiple-section models based on the various approximating schemes

Transconductance (y_{21}) In Fig 4 4(a) and 4 4(b) I compare the phase of y_{21} predicted by the multi-section quasi-static (QS), L and NC models To achieve similar accuracy, the QS model requires about twice as many sections as the L model and the NC model The qualitative reason for the 2 to 1 ratio is as follows

The QS model, the NC model and the L model predict maximum phase shifts of $(n - 1)\pi/2$ $(2n-1)\pi/2$, and $(2n-1)\pi/2$, respectively, where n is the number of subregions of the base region The exact y_{21} has infinite number of poles all of which are negative and real, with no zeros [Lyn67] which means that the maximum phase shift of the exact y_{21} is infinite The ratio of the maximum phase predicted by the QS



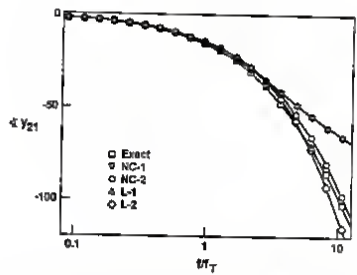


Fig 4.4 The phase of y_{21} predicted by the multi-section quasi-static (QS), L and NC models.
(a) The QS models, where QS-1, QS-2, QS-3 denote one-section two-section and three-section quasi-static models, respectively,
(b) The L and NC models, where NC-1 and NC-2 denote one-section and two-section NC models and L 1 and L 2 denote one-section and two section L models.

andel and by the L (or NC) model is $(2n-1)/(n-1)$ for $n > 1$, thus far a large value of n, the ratio is approximately two.

In Fig. 4.5(a), both the QS and L models correctly predict that $d|y_{21}|/d\omega < 0$ at high frequency. In contrast, the NC model predicts that $d|y_{21}|/d\omega > 0$ at high frequencies (Fig. 4.5b). The significance of $d|y_{21}|/d\omega > 0$ is explored in Chapter Six.

Short-Circuit Current Gain (h_{21}). I compare in Fig. 4.6a and 4.6b the phase of h_{21} vs normalized frequency. The phase of h_{21} predicted by the L model compares favorably with the exact solution (Fig. 4.6a, 4.6b). The NC model, not shown, has a similar accuracy. Note that the QS model, the NC model and the L model predict maximum phase of $\pi/2$, π, and $n\pi$, respectively, where n is the number of sections of the base region. The L model and the NC model have similar accuracy for $|h_{21}|$.

Input Admittance (y_{11}). Comparing Fig. 4.7 with Fig. 3.6a, I observe that the two-section QS model, and the one-section NC and L model predict the phase of y_{11} with similar accuracy.

4.3.2 Comparison with Experiments

Wu et. al. [Wu87] assessed the adequacy of the non quasi-static models by comparing experimental results of h_{11} and h_{21} vs ω with the model predictions. The experiments made use of a large transistor (the unity-gain frequency f_T is about 10 MHz) for which non-quasi-static effects became important at a low frequency. The models used for comparison with measured data are the multi-section L, NC and QS models. For h_{21}, the experiments demonstrated that

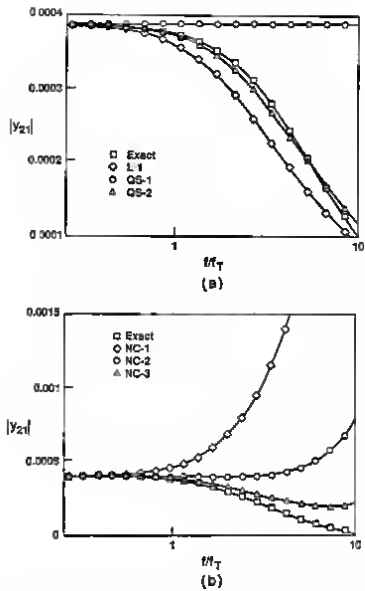


Fig. 4.5 The magnitude of y_{21} predicted by the multi-section QS, L and NC models. See Fig. 4.4 for legends of the curves
(a) The QS and L models,
(b) The NC models

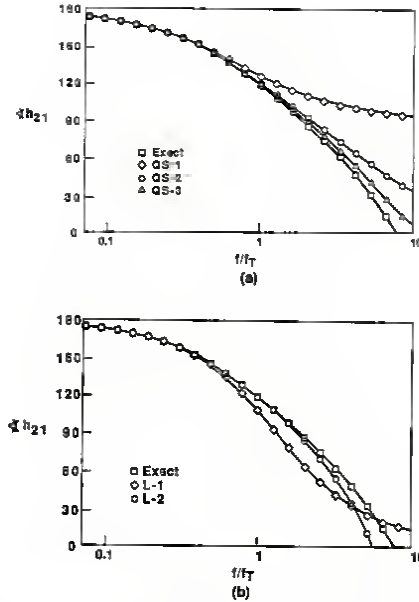


Fig 4 6 The phase of h_{21} predicted by the multi-section QS, and L models See Fig 4 4 for legends of the curves
(a) The QS models,
(b) The L models

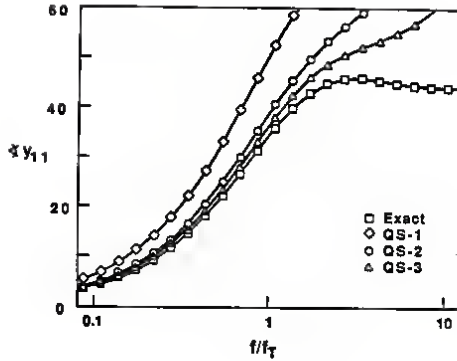


Fig 4 7 The phase of y_{11} predicted by the multi-section QS models See Fig 4 4 for legends of the curves

(1) for frequency below the reciprocal of the base transit time, most of the models agree well with the experiment data except the one-section QS model

(2) a three section L model having six circuit elements for the base region for a transistor in the common-emitter configuration gives a phase response in agreement with experiment for frequency well above the reciprocal of the base transit time. Attainment of this accuracy using quasi-static approximations, which are conventional, would require a model of many more sections

To accurately simulate h_{11}, two-dimensional non-quasi-static models are required. The main reason is that the distributed nature of the base resistance becomes significant for frequency up to the reciprocal of the base transit time. This point is elaborated and demonstrated in Wu and coworkers [Wu87].

4.4 Summary

Taking the delay associated with the convection current into account, I modified existing quasi-static small-signal models for diodes and transistors. Consistently with requirements for computer-aided design, the new non-quasi-static (NQS) equivalent circuits contain no frequency dependent controlled sources. Their derivation proceeds directly from exact solutions and thus directly from the relevant physics.

In Chapter Three, I derived the non-quasi-static models by imbedding a delay element. In the systematic derivation of the models in this chapter, the delay element, instead of being put in entire

from approximate network representations of the minority carrier continuity equations. The two different approaches yield similar non-quasi-static models.

The approach of this chapter works also for quasineutral emitters if one includes the appropriate physics in a suitable form [Wu85]. It incorporates propagation delay of minority carriers in such a way as to maximize accuracy for a given complexity of circuit model.

CHAPTER FIVE
SYSTEMATIC INCLUSION OF DETAILS OF RESPONSE IN LARGE SIGNAL BIPOLAR TRANSISTOR MODELS THE LARGE SIGNAL L AND NC MODELS

Most existing models for circuit simulation programs, for example the Gummel-Poon model in SPICE2, are based on the quasi-static approximation This approximation requires that the carrier densities in the quasi-neutral layers retain their steady-state space dependence on junction voltages even for time-varying excitations The quasi-static approximation in effect assumes that electrons and holes move with infinite velocity thus maintaining during the transient a steady-state space distribution of the hole and electron densities This approximation misses the details of the transient response associated with the time required for minority carriers to propagate in any direction within a quasineutral region Thus for example, in small-signal frequency analysis, quasi static models fail to predict accurately phase margin and stability of wide-band amplifiers In digital-circuit analysis delay is inaccurately described

In this chapter, I present a systematic approach to include the effect of propagation delay in quasi neutral layers The resulting models are applicable to all doping profiles, to all levels of injection and to saturated and forward active operations I start with discussing previous work An extended and improved two-port approach appears in the remaining part of this chapter

5 1 Introduction

Hamilton Lindholm and Norud [Ham64] described a systematic technique for modelling solid-state devices based on what they call a two-port approach The two port approach involves deriving equations for the base layer that have the same form as those encountered in two-port network theory The independent variables are the excess minority carrier densities at the emitter and collector edges of the base, while the dependent variables are emitter and collector minority-carrier current Because of the distributed nature of the base layer, the two-port parameters of these equations are transcendental functions of complex frequency These parameters are then approximated by rational functions

Lynn, et al [Lyn67] applied the two port approach to the bipolar transistor Their work includes a detailed derivation of the model and measurement of the nonlinear model parameters In Section 5 1 1 I briefly review Lynn's work and indicate areas where I shall improve and extend In Section 5 1 2 I summarize the mathematical derivation of the two-port approach, which is the starting point of our extension and improvement of the earlier work

5 1 1 Lynn et al.

Lynn et al [Lyn67] applied the two-port approach to the quasi-neutral base layer as well as the collector and the substrate of an n/p/n bipolar transistor For simplicity I shall concentrate on the quasi-neutral base region only Major assumptions are (1) quasi-

neutrality (2) an exponential doping profile in the base region (3) one-dimensional current flow and (4) low injection

The derivation starts with the continuity equation and current flow j_n for the excess minority carrier n, in the base region.

$$j_n = q \mu_n n E + q D_n \frac{\partial n}{\partial x} \tag{5.1}$$

Letting $j_p = 0$, Lynn obtained the built-in field, and then combined the continuity equation and the minority carrier current flow equation to get a partial differential equation of n (excess minority carrier):

$$\frac{\partial n}{\partial t} = D_n \frac{\partial^2 n}{\partial x^2} - \frac{q D_n}{X_B} \frac{\partial n}{\partial x} - \frac{n}{\tau_n} \tag{5.2}$$

where τ_n is the minority carrier lifetime

Laplace transformation of Eq. (5.2) yields a second-order ordinary differential equation in position x, in which the complex frequency variable s is treated as a parameter. Solving for n as a function of s, Lynn then related the emitter and collector currents to the excess current densities at the edges of the base layer. The boundary condition at the emitter and collector edges are Eq. (4.20)-(4.26) in Lynn et al. [Lyn67]. The solutions for (5.2) in a two port format appear in Eq. (5.5) in Section 5.1.2.

The elements, $A_{ij}(s)$, in the two port equation are transcendental functions. To obtain more tractable solutions, Lynn approximated the frequency dependence of $A_{ij}(s)$. Specifically, Lynn truncated the Taylor series expansions of $A_{ij}(s)$ to obtain single time-constant approximations

$$A_{ii}(s) \approx a_{ii}(1 + s\,r_{ii}) \tag{5 3}$$

$$A_{ij}(s) \approx \frac{a_{ij}}{1 + s\,r_{ij}}, \qquad i \neq j \tag{5 4}$$

I emphasize that, for simplicity, they expanded $A_{ij}(s)$ about $Z = 0$ instead of $s = 0$. The definition of Z is given in Eq. (5 11). After studying the distribution of the poles and zeros of $A_{ij}(s)$, Lynn approximated A_{12} and A_{21} by single-pole functions and A_{11} and A_{22} by single-zero expressions. The approximate model is given in (4 66)-(4 69) in Lynn et al. [Lyn67]. The model gives a two-pole approximation of the forward and inverse common-base current gains (alpha).

They presented a symbolic representation for the nonlinear transistor model, which is shown in Fig. 5 1. The dependent current source in the symbolic representation is nonlinear and frequency-dependent. Gummel [Gum68] noted the undesirability of frequency-dependent controlled sources in computer-aided analysis, and both he and Lindholm [Lin71] developed models for the large signal behavior that contained no frequency-dependent sources yet displayed some aspects of the experimentally observed frequency dependence.

We shall improve and extend Lynn's work in the following aspects:

(1) To obtain a better accuracy, I improve the truncation by expanding $A_{ij}(s)$ about $s = 0$ instead of $Z = 0$. In Section 5 3 1 1 I discuss the error resulting from expansion about $Z = 0$.

(2) I develop an alternate truncation which results in the NC model. The NC model will be connected to other non-quasi-static models

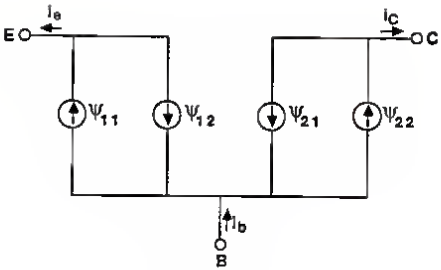


Fig 5 1 Figure 4 4 in Lynn et al [Lyn67] shown in forward-active mode in which $\Psi_{22} = \Psi_{12} = 0$ This is a schematic representation of the transistor nonlinear model for an n/p/n transistor Note that the dependent current sources are nonlinear and frequency dependent The values of Ψ_{11} and Ψ_{21} are given in Eq (4 66) and (4 69) in Lynn and coworkers [Lyn67]

developed by different approaches such as the partitioned-charge based model by Peesum and Veeraraghavan [Pee86] and the transient-integral-charge-control approach by Klose and Wieder [Klo87]. Comparison of recent one-dimensional bipolar transistor models is the subject of Chapter Six.

(3) I develop equivalent circuit representations that have no frequency dependent current sources.

5.1.2 The Two-Part Approach

In this section, I present the mathematical foundation of an approach which includes the effect of propagation delay in quasi-neutral layers. The presentation in this section provides the starting point for our extension and improvement.

This approach is applicable to all doping profiles, to all levels of injection and to saturated and forward active operation. To simplify the derivation at the outset I make the following assumptions: (a) low injection, (b) negligible base recombination, (c) forward active operation, and (d) exponential doping profile in the base layer: that is, $N_A(x)/N_A(0) = \exp(-\eta x/X_B)$, where x=0 defines the emitter edge of the quasi-neutral base. (Many transistors have an approximately exponential doping profile over most of the base layer, although the profile in some transistors leads to a retarding built in electric field near the base edge of the emitter base space-charge layer.) In Section 5.5 I relax these assumptions.

As in Lynn's approach, the derivation begins with the electron continuity equation. Laplace transformation yields a second-order

ordinary differential equation in position x in which the complex frequency variable s is treated as a parameter. The solution of this ordinary differential equation for the transformed excess electron density involves two arbitrary constants. To remove these constants I write equations for the excess electron density and electron current density at the two edges of the quasi-neutral base layer, yielding four equations from which the two constants can be eliminated. The result is a 2X2 matrix equation for the current density at each edge as a function of the excess electron density or the quasi-fermi-potential separation at the edges. This equation resembles those appearing in two-port network theory. As is desirable for models for circuit simulation, no dependence on space variables appears. But the distributed (partial differential-equation) nature of the quasi-neutral base remains imbedded in the transcendental rather than the algebraic functions of s that constitute the elements of the 2X2 matrix.

The elements of the 2X2 matrix are

$$\begin{bmatrix} i_N(0,s) \\ i_N(X_B,s) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} n_E \\ 0 \end{bmatrix}, \tag{5.5}$$

where

$$a_{11}(s) = \frac{q A D_n}{X_B} \left[\frac{\eta}{2} + Z \coth(Z)\right], \tag{5.6}$$

$$a_{22}(s) = \frac{q A D_n}{X_B} \left[\frac{\eta}{2} + Z \coth(Z)\right], \tag{5.7}$$

$$a_{12}(s) = \frac{-q\,A\,D_n\,Z}{X_B \sinh(Z)}\, e^{-\eta/2} \qquad (5\ 8)$$

$$a_{21}(s) = \frac{-q\,A\,D_n\,Z}{X_B \sinh(Z)}\, e^{\eta/2}, \qquad (5\ 9)$$

and $n_B = \frac{n_i^2}{N_A(0)} \left[\exp\left(\frac{V_{BE}}{kT/q}\right) - 1\right],$ (5 10)

in which $Z = \sqrt{(\eta/2)^2 + s\,X_B^2/D_n + (X_B/L_n)^2}$ (5 11)

I assumed negligible base recombination, $L_n = \infty$, hence

$$Z = \sqrt{(\eta/2)^2 + s\,X_B^2/D_n} \qquad (5\ 12)$$

Since I assume forward-active operation the excess electron density at the collector edge of the quasi-neutral base in the second column of (5 5) is zero. Generalization to saturated operation is straightforward for a one-dimensional model.

5 2 Simplification of the Frequency Dependence of the Transistor Model

In Section 5 1 2, I have derived the analytical solutions which form the basis for the development of the improved models. The first step of deriving improved equivalent circuits is to reformulate the two port equation Eq (5 5), to a π equivalent circuit representation.

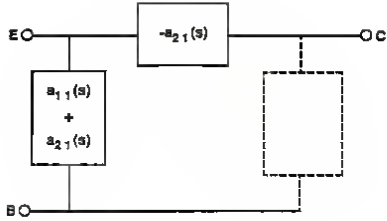


Fig. 5.2 A π-equivalent circuit representation for the two-port equation (5.5). Since we assume forward-active operation the connection (dotted line) between collector and base is an open circuit

which is shown in Fig 5 2 Since I assume forward-active operation the connection between collector and base is open circuited

As indicated earlier, the matrix elements are transcendental functions of s, which reflect the distributed nature of the quasineutral layers embedded in the partial differential equations that describe them To obtain a lumped equivalent-circuit model which corresponds to ordinary differential equations in the independent variable t one must truncate $A_{ij}(s)$ in the complex frequency domain Various truncations are possible Particular truncations produce quasi-static models such as the Gummel Poon model Other truncations produce quasi-static models to which unconventional circuit elements such as inductors or negative capacitors are added These truncations form the basis of the improved models developed in this chapter

5 2 1 Truncation for the Vertical Branch of the π Network

I truncate the matrix elements in the vertical branch of the π-network, which is $a_{11}(s) + a_{21}(s)$ by Taylor expansion about $s = 0$

$$a_{11}(s) + a_{21}(s) = s\, C'_{QS} \tag{5.13}$$

where

$$C'_{QS} = \frac{A\, q\, X_B\, (\eta \quad 1 + e^{-\eta})}{\eta\, (1 - e^{-\eta})} \tag{5.14}$$

If base recombination is included I have an additional conductance in parallel with C_{QS}'

$$a_{11}(s) + a_{21}(s) = G_1' + s\, C_{QS}' \tag{5.15}$$

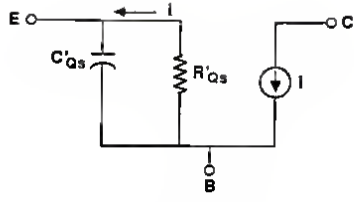
where

$$G_1 = AeD_n/\Delta x \; (e^{\eta/2}\left[\frac{\eta}{2} + \frac{\eta}{2}\coth(\frac{\eta}{2})\right] \quad \frac{\eta/2}{\sinh(\eta/2)}]$$

In the next section three different approximations for the matrix element of the horizontal branch of the π-network, $e_{21}(s)$, result in three models. The approximation for the vertical branch of the π-network, however, remains the same for the three different models. Higher order approximation of the leg of the π network will be discussed in Section 5.7

5.2.2 Truncation for the Horizontal Branch of the π-Network

Several approximations of $e_{21}(s)$, the matrix element in the horizontal portion (her) in the equivalent circuit (Fig. 5.2), are possible and each approximation leads to a different circuit realization, the least accurate of which leads to a quasi-static model. Figure 5.3 shows circuit realizations resulting from three different approximations, Eqs. (5.16), (5.19) through (5.22). Note that the modal variable of the equivalent circuits in Fig. 5.3 is n_g instead of v_{gg} and consequently C' is not a capacitor nor is L' an inductor in the usual sense. I call C and L' "network-like" elements. The equivalent circuits in Fig. 5.3 serve as intermediate steps in deriving the large- and small-signal L and NC equivalent circuits which will be presented in Section 5.3 and 5.4. They also facilitate the comparison of the three different approximations.



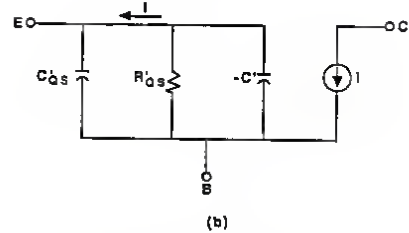
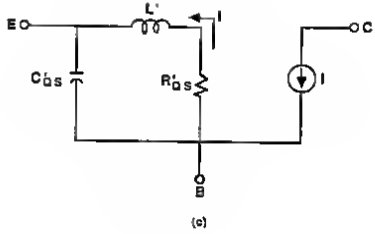


Fig 5 3 Network representation of three equivalent circuits are derived for the quasi-neutral layer of an n/p/n BJT in the forward-active mode The nodal variable is ng for large signal and is $[N_E q/KT]\, v_{be}$ for small-signal The equivalent circuits in this Figure serve as a stepping stone to derive the large- and small signal L and NC equivalent circuits They also facilitate the comparison of the three different approximations (5 16) (5 19) and (5 22)
(a) A quasi-static model (b) A negative capacitance (NC) model (c) An inductive (L) model



(c)

Fig 5 3--continued

The quasi-static model The simplest approximation of $a_{21}(s)$

$$\frac{Z}{\sinh(Z)} = \frac{1}{\left(1 + \frac{Z^2}{6}\right)} \approx 1 \quad (5\ 16)$$

leads to the usual quasi-static equivalent circuit (Fig 5 3a) In the quasi-static equivalent circuit

$$C'_{QS} = \frac{A\, q\, X_E\, (\eta - 1 + e^{-\eta})}{\eta (1 - e^{-\eta})}, \quad (5\ 17)$$

and

$$R'_{QS} = \frac{X_E\, (1 - e^{-\eta})}{A\, q\, D_n\, \eta} \quad (5\ 18)$$

The negative capacitor (NC) model The approximation

$$\frac{Z}{\sinh(Z)} = Z\, \mathrm{csch}(Z) = Z \left(\frac{1}{Z} - \frac{Z}{6} + \frac{7\, Z^3}{360} - \cdots\right) \approx 1 - \frac{Z^2}{6}, \quad (5\ 19)$$

leads to the circuit realization of Fig 5 3(b), a quasi-static equivalent circuit including a negative capacitor (NC) In the NC equivalent circuit (Fig 5 3b)

$$C = -\frac{X_E^2}{D_n} \left(\frac{1 + e^{-\eta}}{\eta (1 - e^{-\eta})} - \frac{2}{\eta^2}\right) \frac{1}{R_{QS}^2} = -\tau_D / R_{QS} \quad (5\ 20)$$

where,

$$\tau_D = \frac{X_E^2}{D_n} \left(\frac{1 + e^{-\eta}}{\eta (1 - e^{-\eta})} - \frac{2}{\eta^2}\right), \quad (5\ 21)$$

in which τ_D is the delay time associated with the minority carrier in the quasi-neutral layer [Che87]

The inductive (L) model The approximation

$$\frac{Z}{\sinh(Z)} = \frac{Z}{\left(Z + \frac{Z^3}{3!} + \frac{Z^5}{5!} + \ \right)} \approx \frac{1}{\left(1 + \frac{Z^2}{6}\right)} \qquad (5.22)$$

replaces the negative capacitor by an inductor (Fig 5 3c) In the L equivalent circuit (Fig 5 3c),

$$L' = \tau_D \theta_{QS} \qquad (5.23)$$

5.3 The Large Signal NC Model

The modal variable of the equivalent circuits in Fig 5 3 is n_E In Section 5 3 and 5 4, I transform the physical equivalent circuits, in which n_E is the modal variable, to large-signal (electrical) equivalent circuits, in which voltage is the modal variable

I start the transformation of the NC model (Fig 5 3b) by rewriting the current in the network like capacitors in Fig 5 3(b) as

$$C_{QS} \frac{dv_{gE}}{dt} = C'_{QS} \frac{dn_E}{dt} = C'_{QS} \frac{n_E}{kT/q} \frac{dv_{gE}}{dt} \qquad (5.24)$$

and

$$C \frac{dv_{gE}}{dt} = C' \frac{dn_E}{dt} = C' \frac{n_E}{kT/q} \frac{dv_{gE}}{dt} \qquad (5.25)$$

I then rewrite the current flowing through the network like resistor

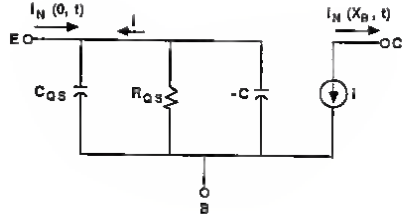


Fig 5 4 The large-signal NC model for the quasi neutral base of an n/p/n BJT in the forward-active mode All elements depend on $v_{BE}(t)$

$$\frac{n_E}{R_{QS}} = \frac{v_{BE}}{(R'_{QS}\, v_{BE}/n_E)} = \frac{v_{BE}}{R_{QS}} \tag{5.26}$$

Thus, I derive the large signal NC equivalent circuit in Fig 5.4 with the nonlinear capacitor

$$C_{QS} = C'_{QS}\, \frac{n_E}{kT/q} ,$$

which is the conventional diffusion capacitance defined by

$$\frac{\partial |q_{QS}|}{\partial v_{BE}} = \frac{A\, q\, \eta\, n_i^2 X_B\, (\eta\, 1+e^{-\eta})}{N_A(0)\, \eta^2\, (1-e^{-\eta})\, kT/q}\, e^{\frac{v_{BE}(t)}{kT/q}} \approx \frac{|i_{QS}|\, \tau_{tr}}{kT/q} , \tag{5.27}$$

where

$$q_{QS}(t) = \int_0^{X_B} q\, n_{QS}(x,t)\, dx = i_{QS}(t)\, \tau_{tr} \tag{5.28}$$

in which

$$|i_{QS}(t)| = \frac{A\, q\, D_n\, n_i^2\, \eta}{N_A(0)\, X_B\, (1 - e^{-\eta})} \left[\exp\left(\frac{v_{BE}(t)}{kT/q}\right) - 1\right] \tag{5.29}$$

and

$$\tau_{tr} = \frac{X_B^2}{D_n}\, \frac{\eta - 1 + e^{-\eta}}{\eta^2} \tag{5.30}$$

For an n/p/n transistor, $i_{QS} < 0$ and $q_{QS} < 0$. The other two nonlinear elements in Fig 5.4 are

$$-C = -\tau_D\, |i_{QS}|/(kT/q) , \tag{5.31}$$

and

$$R_{QS} = \frac{R'_{QS}\, v_{ZE}}{n_E} = \frac{v_{ZE}}{|I_{QS}|} \qquad (5\ 32)$$

For a uniformly doped base layer,

$$|I_{QS}| = \frac{A q\, D_n n_i^2}{N_A(0)\, X_B} \left[\exp\left(\frac{v_{ZE}}{kT/q}\right) - 1\right] \qquad (5\ 33)$$

$$\tau_{tr} = X_B^2/2D_n \quad \text{and} \quad \tau_D = \tau_{tr}/3 \qquad (5\ 34)$$

5 4 The Large Signal L Model

The series connection of L' with R_E in Fig 5 3(b) precludes directly transforming n_E to v_{BE} as done for the large-signal NC model in section 5 4 1. I discuss several ways to do this transformation. Among these methods, I present a better transformation in Section 5 4 2.

5 4 1 Options for Deriving the Large Signal L Model

One option would be to use the physical equivalent circuit with n_E as the node variable instead of using an electrical equivalent circuit. The advantage of this approach is that I avoid the sometimes complicated transformation from the physical equivalent circuit to the electrical equivalent circuit. Unfortunately, circuit simulation programs do not use n_E as a node variable.

A second option is to introduce additional nodes, for example, n' and v' in Fig. 5 5(a) and Fig. 5 5(b), respectively. These additional nodes enable the direct transformation by equating the current flow

through L' and R' in Fig 5 5(a) to the current flow through L and R in Fig 5 5(b)

$$\frac{1}{L'}\left[\int(n_E - n')dt\right] = n'/R' = i = \frac{1}{L}\int(v_E - v')\,dt = v/R$$

The disadvantage of this option is that the additional node reduces the computational efficiency of the computer simulation. Besides, the elements L and R derived in this way are functions of n' which is undesirable because n' is not known.

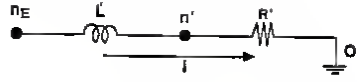
A third alternative would be to take the Laplace transformation of the model variable n_E and then apply the Laplace operation to derive the large signal model as is done in the treatment of the small signal models. The problem is that no Laplace transformation of $\exp(q\,v_{BE}(t)/kT)$ in $n_E(t)$ exists.

A final option is to solve for the current flowing through R' and L' in Fig 5 5(a) and use the solution for the model. This approach results in two current sources in parallel, one accounting for the convection current and the other one accounting for the propagation delay. This model is computationally efficient because no nodes are added to the equivalent circuit. This approach is discussed in detail in the next section.

5 4 2 Derivation of the Large-Signal L Model

I shall derive the large-signal L model by solving for the current flowing through L' and R' in Fig 5 3(b):

$$L'\frac{di}{dt} + R'i = n_E \qquad (5\ 35)$$



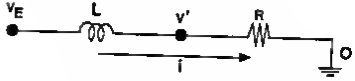


Fig 5 5 Two equivalent circuits with different nodal variable (a) The physical equivalent circuit, in which the nodal variable is the excess minority carrier density. (b) The electrical equivalent circuit, in which voltage is the nodal variable. The transformation from (a) to (b) is not straightforward

which as shown in Rose [Ros74] has as its solution:

$$i(t) = e^{-t/\tau_D}\left[\int e^{t/\tau_D}\frac{n_E(t)}{L'}\,dt + i(0) - F(0)\right] \qquad (5.36)$$

where $F(t) = \int e^{t/\tau_D}\frac{n_E(t)}{L'}\,dt = \int_0^t e^{t/\tau_D}\frac{n_E(t)}{L'}\,dt + F(0)$ (5.37)

After some mathematical manipulation to be shown later I rewrite Eq. (5.36) as Eq. (5.38)

$$i(t) = |I_{QS}(t)| + \frac{d}{dt}[q_{G(L)}] \qquad (5.38)$$

where

$$|I_{QS}(t)| = \frac{n_E(t)}{\tau} , \qquad (5.39)$$

and

$$q_{G(L)} = \frac{e^{-t/\tau_D}}{\tau'_{QS}}\int_0^t e^{\xi/\tau_D} n_E(\xi)\,d\xi - i(0)\tau_D e^{-t/\tau_D} \qquad (5.40)$$

The current expression in Eq. (5.38) which is in terms of charge-based representations consists of two current components. The first component is the convection or conduction current, which is equivalent to the Gummel formulation [Gum68]. The second term is a charging current, which accounts for the non-quasi-static effect. The large signal L model is shown in Fig. 5.6

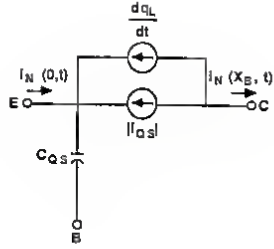


Fig 5.6 The large signal L model for the quasi neutral base of an n/p/n BJT in the forward active model All elements depend on $v_{BE}(t)$

I now give a brief description of the derivation of Eq. (5.38). For details of the mathematical manipulation, see Appendix A. The first step to derive Eq. (5.38) is to integrate the first term in the right-hand side of Eq. (5.36) by part:

$$e^{\frac{-t}{\tau_D}} \int e^{\frac{t}{\tau_D}} \frac{n_g(t)}{L'} dt = \frac{n_g(t)}{L'/\tau_D} - \frac{e^{-t/\tau_D}}{L'/\tau_D} \int e^{\frac{t}{\tau_D}} \frac{dn_g(t)}{dt} dt \qquad (5.41)$$

Substituting (5.41) into (5.36) and replacing L'/τ_D by R_{QS}' yields

$$i(t) = \frac{n_g(t)}{R_{QS}'} - \frac{e^{-t/\tau_D}}{R_{QS}'} \int e^{\frac{t}{\tau_D}} \frac{dn_g(t)}{dt} dt + e^{\frac{-t}{\tau_D}} [i(0) - F(0)] \qquad (5.42)$$

The second step is to move d/dt out of the integral in the second term of the right-hand side of Eq. (5.42). As a result, Eq. (5.42) becomes

$$i(t) = \frac{n_g(t)}{R_{QS}'} - \frac{d}{dt}\left[\frac{e^{-t/\tau_D}}{R_{QS}'} \int e^{\frac{t}{\tau_D}} n_g(t)\, dt\right] + e^{\frac{-t}{\tau_D}} [i(0) - F(0)] \qquad (5.43)$$

The third step is to replace the indefinite integral in Eq. (5.43) by a definite integral by combining the second term with F(0) in the right-hand side of Eq. (5.43). The result is Eq. (5.38), which completes the derivation.

5.5 Comparison of the Large-Signal L and NC Models

To summarize the derivations of the large-signal L model and the large-signal NC model in Section 5.3 and 5.4, I compare $i(X_B,t)$ for the large-signal L and NC models, (5.44) and (5.45), respectively.

$$i_{N(L)}(X_B,t) = i(t) - |i_{QS}(t)| + dq_{C(L)}/dt \qquad (5.44)$$

$$i_{N(NC)}(X_B,t) = |i_{QS}(t)| + dq_{C(NC)}/dt \qquad (5.45)$$

where

$$q_{C(NC)} = i_{QS}\,\tau_D = \frac{q\,n_E(t)\,D_n}{X_B}\,\tau_D < 0 \qquad (5.46)$$

$$q_{C(L)} = -\frac{e^{-t/\tau_D}}{R'_{QS}} \int_0^t e^{\xi/\tau_D} n_E(\xi)\,d\xi - i(0)\tau_D\, e^{-t/\tau_D} \qquad (5.47)$$

Both $i(X_B,t)$ for the large-signal L and the large signal NC model consist of two components. The first component is the convection current or conduction current which is equivalent to the Gummel formulation [Gum70a]. The second component is a charging current that account for the non-quasi-static effects. The difference between the L model and the NC model is in handling of the charging components. In comparing the two models, two questions arise. First, what is the significance of $i(0)$ in (5.47). Second, how does the integration in Eq. (5.47) effect the computer simulation speed? The first question will be answered in Section 5.5.1. In Section 5.5.2 I briefly compare the computational efficiency of the two models.

5.5.1 Initial Condition $i(0)$

In the steady state $n_E(t)$ is a constant and Eq. (5.47) can be rewritten as (5.48) by moving n_E out of the integral.

$$q_{C(L)} = \frac{n_E}{R_{QS}} (1 - e^{-t/\tau_D}) \tau_D \quad i(0) \tau_D e^{-t/\tau_D} \tag{5.48}$$

Since $i(0) = n_E(0)/R_{QS}$

$$q_{C(L)} = n_E/R_{QS}'$$

which consequently makes $dq_{C(L)}/dt = 0$. In other words the presence of $i(0)$ enables $q_{C(L)}$ to be time independent in the steady state.

For $q_{C(NC)}$ in the NC model the initial current is implicitly accounted for. This is verified by letting $t = 0$ in Eq. (5.48),

$$q_{C(NC)}(t=0) = \frac{q\, n_E(0)\, D_n}{X_E} \tau_D = i_{QS}(t=0)\, \tau_D, \tag{5.49}$$

where $i_{QS}(t=0)$ is the initial current. Thus (5.46) is not a function of time at steady state and the presence of $i(0)$ is not required.

5.5.2 Computational Efficiency

I shall briefly compare the computational efficiency of the large-signal L and NC models. For details of the computer implementation of the large-signal L model and the large signal NC model, see Chapter Seven.

To evaluate the L and NC models from the viewpoint of computer simulation, I compare the Eq. (5.44) with Eq. (5.45). From (5.44) and (5.45) the difference in simulation speed between the L and the NC models will be determined by the calculation of $dq_{C(L)}/dt$ and $dq_{C(NC)}/dt$. For the L model the CPU time required for circuit

analysis is comparable to that required to the NC model. This is true because instead of evaluating the integral in (5.47) from 0 to t by numerical integration, I only need to determine the integral from t-Δt to t, which requires little CPU time. Thus, at time point t_{n+1}, I rewrite the integral portion of (5.47) as (5.50)

$$\int_0^{t_{n+1}} e^{\zeta/\tau_D} n_g(\zeta)d\zeta = \int_0^{t_n} e^{\zeta/\tau_D} n_g(\zeta)d\zeta + \int_{t_n}^{t_{n+1}} e^{\zeta/\tau_D} n_g(\zeta)d\zeta \qquad (5.50)$$

The first term on the right-hand-side of (5.50) is known by keeping track of solutions from previous time points. The second integral on the right-hand-side of (5.50) is approximated by the trapezoidal integration method

$$\int_{t_n}^{t_{n+1}} e^{\zeta/\tau_D} n_g(\zeta)d\zeta = \frac{\Delta t}{2}\left[e^{t_n/\tau_D} n_g(t_n) + e^{t_{n+1}/\tau_D} n_g(t_{n+1})\right], \qquad (5.51)$$

where $\Delta t = t_{n+1} - t_n$. Since Δt is kept small to maintain an acceptable local truncation error of the numerical integration method, (5.51) is accurate.

5.6 Small-Signal NC and L Models

I start the small signal derivation with the approximation in (5.52),

$$e^{V_{BE}/(kT/q)} = e^{(V_{gg}+v_{be})/(kT/q)} \approx e^{V_{gg}/(kT/q)}\left[1 + \frac{v_{be}}{kT/q}\right]. \qquad (5.52)$$

where V_{gg} and v_{be} are the quiescent and the small-signal quantities respectively. Equation (5.52) indicates that the matrix equation (5.5) can be separated into dc and small-signal components. Thus (5.5) would be the matrix equation for the small signal excitation provided that I replace i_N by i_n and replace n_g by $N_g\, v_{be}/(kT/q)$ where N_g is (5.10) with V_{gg} instead of v_{gg}. Using approximations similar to those of (5.16), (5.19) and (5.22), I derive the small signal circuit representations in Fig. 5.3 with the identical network-like elements, (5.17) through (5.23), provided that the model variable is the weighted potential, $[N_g/(kT/q)]\, v_{be}$ instead of n_g for the large-signal excitation. Unlike the previous derivation for large-signal models, however, I can put the weighted potential directly into the network-like elements and obtain small signal models in which v_{be} is the model variable. The resulting circuit configuration of the small-signal NC model resembles Fig. 5.3(a) but has linear circuit elements

$$R_{qe}^{ss} = \frac{kT/q}{I_{QS}} \qquad C_{qe}^{ss} = \frac{I_{QS}\, \tau_{tr}}{kT/q}\ , \text{ and } -C^{ss} = -\frac{\tau_D\, I_{QS}}{kT/q} \tag{5.53}$$

where ss stands for small signal and I_{QS} is (5.29) with V_{gg} replacing v_{gg}. The small-signal L model is similar to Fig. 5.3(b) with

$$L^{ss} = \tau_D\, R_{qa}^{ss} \tag{5.54}$$

5.7 Higher-Order Approximations

In this section, I discuss higher order truncation of the matrix elements which results in more complicated equivalent circuits

Recall that in Section 5 2 1 I approximated the matrix elements in the vertical branch of the π equivalent circuit of the two-port network (Fig 5 2) by Eq (5 13). Two higher-order approximations of the matrix elements in the vertical branch of the π equivalent circuit are presented next. For simplicity, I assume a uniformly doped base region. Extension to an exponentially doped base is straightforward.

5 7 1 Seitchik's Model

One model derivable from a higher-order approximation was proposed by Seitchik and coworkers [Sei57], which includes a capacitor and a resistor in series to account for the delay time associated with the stored charge in the quasi-neutral base region. I now derive Seitchik's model using our systematic approach. For a uniformly doped base the vertical branch of the π-network are

$$z_{11}(s) + z_{12}(s) = \frac{A\, q\, D_n}{X_B}\left[Z \coth(Z) - \frac{Z}{\sinh(Z)}\right]$$

$$= \frac{A\, q\, D_n}{X_B}\left[\frac{\frac{Z^2}{2!} + \frac{Z^4}{4!} + \cdots}{1 + \frac{Z^2}{3!} + \cdots}\right] \qquad (5\text{-}55)$$

where $Z = (s\frac{X_B^2}{D_n})^{1/2}$ for a uniformly doped base

Recall that the approximation discussed in Section 5 2 1, Eq (5 13), is obtained by retaining the first term in the numerator, $Z^2/2!$ and the first term in the denominator. Thus

$$\frac{A\,q\,D_n}{X_E}\,\frac{Z^2}{2!\,i} = \frac{A\,q\,X_E}{2}\,\varepsilon = \varepsilon\,C'_{QS} \qquad (5\ 56)$$

To obtain one level of approximation above (5 56) I keep the s^1 dependence in the denominator of Eq (5 55) Thus

$$\text{Eq (5 55)} = \frac{\varepsilon\,A\,q\,X_E}{2}\,\frac{1}{1+\varepsilon\,X_E^2/12D_n} = \frac{\varepsilon\,C'_{QS}}{1+\varepsilon C'_{QS}\,R'_{SD}} \qquad (5\ 57)$$

where $R_{SD}' = X_E/(6\,A\,\varepsilon\,D_n)$ Eq (5 57) can be modeled by a resister-like element, R_{SD}', in series with C_{QS}' as shown in Fig 5 7(b) Note that the modal variable in Fig 5 7 is n_E This improvement has been suggested by Seitchik et al [Sei67] to include the delay in the stored charge of the quasi neutral layer during transient The difference between Seitchik's work and the derivation just described is that I use a systematic approach instead of the derivation of Seitchik and coworkers [Sei67] This model predicts input impedence more accurately, as discussed in Chapter Three

5 7 2 Higher-Level Approximation

The next higher level approximation of Eq (5 57) can be obtained by retaining the s^2 dependence in the numerator of (5 55)

$$\text{Eq (5 55)} = \frac{\varepsilon\,\frac{A\,q\,X_E}{2} + \varepsilon^2\,\frac{A\,q\,X_E^3}{24\,D_n}}{1+\varepsilon\,X_E^2/12D_n}$$

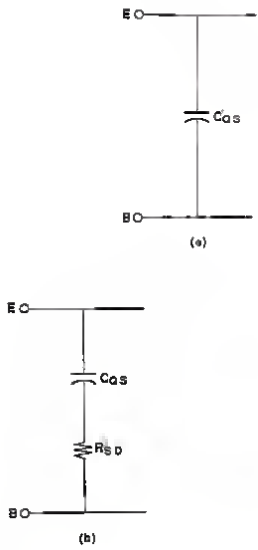


Fig. 5.7 Three equivalent circuits resulting from different levels of approximation on the matrix elements in the leg of the π-network representation of the two-port network
(a) The first level, Eq. (5.53); (b) The second level, Eq. (5.54); (c) The third level, Eq. (5.55)

$$= \frac{s\,C_{QS}' + s^2 \frac{C_{QS}' C_2'}{G_S}}{1 + s \frac{C_{QS}' + C_2'}{G_S}} \qquad (5\ 58)$$

where $C_2' = A \cdot q\, X_B/2$ $G_S' = 6\, A \cdot q\, D_n/X_B$ This approximation is implemented in equivalent circuit form as shown in Fig 5 7(c)

As in the equivalent circuits shown in Fig 5.3, whose modal variable is excess minority carrier density instead of voltage, the equivalent circuits shown in Fig 5 7 are not electrical equivalent circuits The transformation of the equivalent circuits in Fig 5 7 to electrical equivalent circuits is very complicated and will not be pursued here.

5 8 Extensions of the L and NC Models

In deriving the L and NC models I assumed (a) forward-active, (b) low injection (c) no two-dimensional effects, such as emitter crowding (d) low enough doping concentration that energy-gap narrowing is negligible, (e) an exponentially doped base region and (f) negligible effects from base width modulation base resistance, and base push-out The purpose of this Section is to relax assumptions (a) through (e) by inserting more physics into the model

Wu et al [WuE7] has relaxed assumption (b) to model high and moderate injection assumption (c) to account for the ac and dc emitter crowding effects, assumption (d) to extend the non-quasi static models to quasi-neutral emitter and (e) to accommodate arbitrary doping

profiles. I shall briefly indicate how these assumptions are relaxed. For details, see Wu and coworkers [Wu87].

Recall that our systematic approach starts with Laplace transformation of the continuity equations with respect to time. By this procedure the complex frequency variable, $s = \sigma + j\omega$, replaces the independent variable time t. In a one-dimensional model Laplace transformation with respect to time directly reduces partial differential equations. In a two-dimensional model the reduction of a two-dimensional differential equations to one-dimensional ordinary differential equations requires the additional use of the product method or separation of variables method familiar from the theory of differential equations. The analytical solutions form the basis for development of the models.

Extension (a), saturated mode of operation. The extension to saturated mode of operation is done by superposing the results from the forward-active and reverse-active operations.

Extension (b), two-dimensional effects. Two-dimensional emitter transient (or ac) crowding is most important for two cases: when high injection occurs in the base and when the high frequency portion of the Fourier spectrum of the input excitation is significant. The systematic approach to account for crowding, developed by Wu et al. [Wu87], involves integrating the two-dimensional continuity equation in the direction parallel to the underside of the emitter-base junction. Coupling between the continuity and current density equations makes iterative solutions necessary.

Extension (c), heavy doping effects The inclusion of heavy doping effects results in a non-constant coefficient in the ordinary differential equation which can be solved by an integration and iteration mathematical scheme See Wu and Lindholm [Wu89] for details

Extension (d), high and moderate levels of injection To account for the non-quasi-static effects in high injection level Wu and Lindholm [Wu89] solves the ambipolar transport equation Moderate levels of injection are modeled by a numerical fitting technique between low and high injection, yielding a general model for all injection levels This approach extends a method used by de Graaff and Kloosterman [Deg85]

Extension (e), arbitrary doping profile I account for the arbitrary doping profile by approximating the arbitrary doping profile by piecewise constant profile I then divide the quasi-neutral base layer into subregions that define segments of the piecewise-linear curve Applying the systematic method to each subregion results in a complete model accounts for the arbitrary doping profile Details are in Wu and Lindholm [Wu89]

5 9 Discussion

Next I discuss two other approaches related to the expansions of the matrix elements, $A_{ij}(\epsilon)$ In Section 5 9 1 I compare the expansion of Taylor series about $\epsilon = 0$ with the expansion of Taylor series about $Z = 0$ In Section 5 9 2, I discuss the expansion of the transcendental function by Taylor series versus by Pade expansion

> 9 1 Taylor Expansion About Z = 0

Lynn et al [Lyn67] expanded the matrix elements associated with the two-port equation by Taylor series about Z = 0 I expanded the matrix elements by Taylor series about s = 0 to ensure correct DC characteristics Here I compare the two approaches

To assess the two expansions I compare the steady-state collector current predicted by Lynn's model with that given by the L (or NC) model for an exponentially doped base The standard for comparison is the exact steady-state collector current, which is obtained by solving steady state continuity equation together with the electron current density equation:

$$I_C = \frac{A\, q\, D_n\, n_i^2\, \eta}{N_A(0)\, X_B\, (1 - e^{-\eta})} \left[\exp\left(\frac{V_{BE}}{kT/q}\right) - 1\right], \qquad (5\ 59)$$

The OC collector current predicted by Lynn is given in (4 71) in Lynn et al [Lyn67] which is rewritten as Eq (5 60) by letting $X_B/\tau_n \to 0$ because of the assumed negligible base recombination

$$I_C = a_{21}\left(e^{\frac{q\, V_{BE}}{kT}} - 1\right) \qquad (5\ 60)$$

where

$$a_{21} = \frac{A\, q\, D_n\, n_i^2\, e^{\eta/2}}{N_A(0)\, X_B\, [1 + \eta^2/24]}$$

The steady-state I_C predicted by the NC model and the L model is given in Eq (5 29) which is equivalent to the exact steady state

collector current (5.59). This equivalence proves that Taylor expansion about $s = 0$ gives correct DC characteristics.

Comparing (5.59) with (5.60), I note that (5.60) is exact for $\eta = 0$ (uniformly doped base). This is because for $\eta = 0$, Z reduces to $(s X_B/D_n)^{1/2}$ and consequently the expansion about $Z = 0$ is equivalent to the expansion about $s = 0$. For $\eta > 0$, however, Eq. (5.60) overestimates I_C. As an example, when $\eta = 8$, Eq. (5.60) overestimate I_C by 86% (Fig. 5.8).

I next discuss the effect of Taylor expansion about $Z = 0$ on τ_D, which is the coefficient of the s^1 term in $A_{21}(s)$ or $i_C(x,s)$. Note that $i_C(X_B,s) = n_E A_{21}(s)$. The standard for comparison in the exact solution for τ_D, which is obtained from the exact $i_C(X_B,s)$. For an exponentially doped base layer, the delay time predicted by Lynn, $\tau_{D(Lynn)}$, is given by (4.78) in Lynn et al. [Lyn67], which is repeated in (5.61) with the emphasis on the expansion about $Z = 0$.

$$A_{21}(s) = A_{21}(Z=0) + Z \left.\frac{dA_{21}(Z)}{dZ}\right|_{Z=0} + \cdots \tag{5.61}$$

where

$$\tau_{D(Lynn)} = \frac{\left.\frac{dA_{21}(Z)}{dZ}\right|_{Z=0}}{A_{21}(Z=0)} = \frac{X_B^2}{6 D_n} \frac{1}{(1 + \frac{1}{6} Z_0)} \tag{5.62}$$

in which, for negligible recombination

$$Z_0 = (\frac{\eta}{2})^2 + (\frac{X_B}{L_n})^2 \rightarrow (\frac{\eta}{2})^2 \tag{5.63}$$

Similar results for τ_D are also obtained in Jung [Jun86].

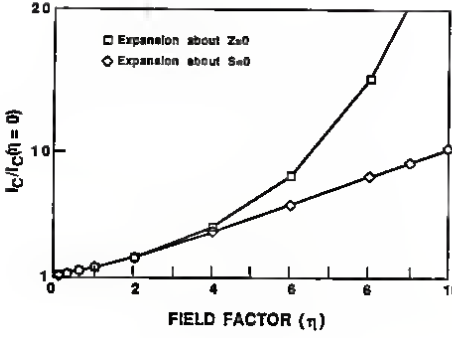


Fig 5 8 The steady state collector current I_C, predicted by Lynn et al [Lyn67] and by the L and NC models The collector current predicted by the L and NC model is equivalent to the exact I_C Lynn et al overestimates i_C by 86% for $\eta = 8$

Taylor expansion about s = 0 results in the following formulation for τ_D

$$A_{21}(s) = A_{21}(s=0) + s \left.\frac{dA_{21}(s)}{ds}\right|_{s=0} + \cdots \tag{5.64}$$

$$\tau_D = \frac{\left.\frac{dA_{21}(s)}{ds}\right|_{s=0}}{A_{21}(s=0)} = \frac{X_B^2}{D_n} \left(\frac{\sum_{n=1}^{\infty} \frac{n z_0^{n-1}}{(2n+1)!}}{\sum_{m=0}^{\infty} \frac{z_0^m}{(2m+1)!}} \right) \tag{5.65}$$

Note that if I let n = 1 and m = 1 Eq (5.65) reduces to Eq. (5.62) With some mathematical manipulation, the summation in (5.65) can also be written as

$$\tau_D = \frac{X_B^2}{D_n} \left(\frac{1 + e^{-2\sqrt{z_0}}}{2\sqrt{z_0}(1 - e^{-2\sqrt{z_0}})} - \frac{1}{2 z_0} \right) \tag{5.66}$$

Substituting (5.63) into (5.66) yields

$$\tau_D = \frac{X_B^2}{D_n} \left(\frac{1 + e^{-\eta}}{\eta (1 - e^{-\eta})} - \frac{2}{\eta^2} \right) \tag{5.67}$$

Equation (5.67) is equivalent to the exact solution (See Eq. 5.22) Equation (5.62) underestimates the delay time by 51% for $\eta = 8$ (See Fig. 5.9).

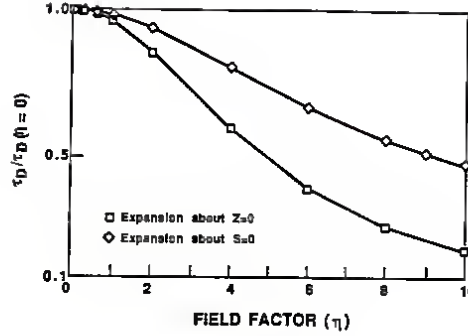


Fig 5 9 The delay time τ_D predicted by Lynn et al [Lyn67] and by the L and NC models The delay time predicted by the L and NC model is equivalent to the exact delay time The delay time predicted by Lynn et al underestimates by 51% for $\eta = 8$

5.9.2 Padé Expansion

Engl and Kleanstelidis [Eng69] employed Padé expansions rather than Taylor expansions to obtain rational functions for the two-port parameters. Padé expansions expand the exponential function by (5.68)

$$\exp(x) = P_m(x)/Q_n(x) \tag{5.68}$$

where P and Q are polynomials of mth and nth order, respectively. By systematically choosing the degree of the polynomials, Engl and Kleanstellidis obtain a model which, in its lowest complexity, has the same form as the Ebers-Moll model [Ebe54]. If the complexity is increased by one degree, the resulting model has the same form as the Nerud-Meyer model [Lyn67], which predicts a two-pole alpha (current gain), an increase of complexity of one more degree yields a model having a two-pole, single zero alpha.

Our discussion will concentrate on their model with the highest degree of complexity, in which a LHP (left-half-plane) zero is used to approximate the $a_{11}(s)$ and $a_{22}(s)$, a LHP pole and a RHP (right-half-plane) zero are used to approximate $a_{21}(s)$ and $a_{12}(s)$. Since the current gain is $a_{21}(s)/a_{11}(s)$, the model predicts a two-pole, single zero alpha.

Engl approximated $a_{21}(s)$ by a RHP zero and a LHP pole. To properly evaluate the accuracy of their approach, I should compare Engl's model with a model that approximates $a_{21}(s)$ (or a_{12}) by a two LHP poles, which can obtained by retaining the s^2 term in Taylor expansion of $a_{21}(s)$, Eq. (5.8)

Since $s_{21}(s)$ and $s_{12}(s)$ have an infinite number of poles and no zeros [Lyn67], it is more accurate to approximate $s_{21}(s)$ (or s_{12}) by two LHP poles than by a RHP pole and a RHP zero. Note that a RHP zero predicts same phase shift as an equivalent LHP pole. As frequency increases, however, the RHP zero predicts an increase in magnitude, which is incorrect. In addition, it is difficult to implement the higher-order model in an equivalent circuit format.

Because of the incorrect trend in predicting the magnitude of $s_{21}(s)$ and the difficulty in implementing an equivalent circuit, I conclude that there is no advantage of the Engl's approach [Eng69] over our systematic approach.

5.10 Summary

In this chapter, I have accomplished the following.

1. I improved Lynn's work by expanding $A_{ij}(s)$ about $s = 0$ instead of $Z = 0$. For $\eta = 8$, Lynn's model overestimates 1_C by 86% and underestimates the delay time, τ_D, by 51%.
2. I included the non-quasi-static effect of propagation delay directly from the continuity equations and associated physics. The resulting large- and small-signal L and NC equivalent circuits, extended by Wu et al [Wu67], are applicable to all doping profiles, to all levels of injection, to transient or ac emitter crowding in the quasineutral base, emitter and collector layers and for forward-active, reverse active and saturated operation.
3. I used a systematic approach, instead of the less systematic derivation of Seitchik et al [Sei67], to derive Seitchik's model

4. I showed that the approach proposed by Engl and Kleustelidis [Eng69] has an advantage over our systematic approach.

Accounting for the propagation delay as described in this chapter provides a basis for more accurate circuit simulation. Despite the detailed physics included, the models are compact enough for use in circuit simulators. They also reduce the problem of parameter extraction because the additional physics imbedded in the model reduces the number of parameters that need to be extracted.

These models have nearly a traditional equivalent circuit form. They consist of passive circuit elements and controlled sources; consistently with the exact model for the basic semiconductor device (Shockley) equations put forward by Sah [Sah70]. The models are not exactly traditional equivalent circuits because unconventional elements, negative capacitances or inductances, appear in them. These elements serve to include non quasi static effects while restraining the overall model complexity as measured by the number of circuit element or number of poles and zeroes.

CHAPTER SIX
COMPARISON AND EXTENSION OF RECENT ONE-DIMENSIONAL BIPOLAR TRANSISTOR MODELS

6.1 Introduction

In this chapter, I compare and evaluate four models that account for non-quasi static effects in the quasi neutral base layer. These models are the partitioned charge-based (PCB) model by Fossum and Veeraraghavan [Fos86], the transient integral charge control model (TICC) by Klosa and Wieder [Kle87], and the negative-capacitance (NC) and the inductive (L) equivalent circuits derived by truncation in the s (complex frequency) domain. A derivation of NC and L equivalent circuits appears in Chapter Five.

In Section 6.2, I briefly sketch the three approaches. More details about the negative-capacitance (NC) and inductive (L) models appeared in Chapter Five. Section 6.3 demonstrates the equivalence among the PCB, TICC and NC models under certain restrictions. In Section 6.4.1 I show that none of these three models correctly gives the trend that the magnitude of the small signal transconductance follows at high frequencies. In Section 6.5, I indicate that the L model does, and thus is more accurate.

6 2 Brief Sketch of Non Quasi Static Models

We assume throughout this chapter an n/p/n bipolar transistor We focus on the quasi neutral base layer and neglect base recombination

6 2 1 Partitioned-Charge-Based Model (PCB)

Fossum and Veeraraghavan [FosS6] included non-quasi static effects in a quasi-neutral base layer in low injection Assuming the forward active mode of operation and negligible base recombination they integrated the electron continuity equation with respect to position x twice inserting the quasi-static approximation after the second integration This approach contrasts with the conventional charge-control method in which the quasi static approximation is inserted after a single integration The resulting partitioned-charge-based (PCB) model (Fig 6 1) applied to the analysis of an emitter-coupled logic (ECL) gate showed details in the input output characteristic not obtainable from traditional charge-control modeling [FosS6] The derivation of the PCB model involves partitioning the total injected base charge into two portions one associated with the emitter and another associated with the collector These authors indicated that the optimum partitioning is not well defined if the base has a built-in electric field Thus the PCB model as derived in Fossum and Veeraraghavan [FosS6], is accurate and defined completely only for a uniformly doped base layer in low injection

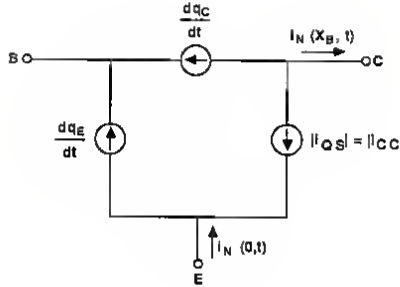


Fig 6 1 Network representation of the PCE model derived for the quasi-neutral base of an n/p/n BJT in the forward active mode In this large-signal model, all elements depend on $v_{BE}(t)$.

6.2.2 Transient-Integral-Charge-Control Relation (TICC)

Klose and Wieder [Klo87] developed the transient-integral-charge-control (TICC) relation which resembles Gummel's formulation [Gum70a] yet is extended by one additional term. The TICC relation provides a closed-form derivation for both the dc and the transient behavior and predicts the partition of base charge during transients.

In their derivation, they integrated the electron current density equation with respect to position. At first, the two limits of the integration were chosen arbitrary. Klose and Wieder then performed a second integration via partial integration rules and inserted the electron continuity equation, setting the two arbitrary integration limits as the base and collector edges of the quasi neutral base layer. The resulting two current expressions, called the TICC relation, appear as equation (4) in Klose and Wider [Klo87]. Each current consists of two components. The first component is the convection or conduction current which is equivalent to the Gummel formulation [Gum70a] extended to include energy gap narrowing and Fermi statistics through the substitution $n_i \rightarrow n_{ie}$ [Kle87]. The second component is a charging current, which is a function of the excess electron density multiplied by a weight function. (In Klose and Wieder [Klo87], it is stated to be the electron density multiplied by a weight function. From context, apparently it is the excess electron density.) To determine the weight function, however, requires the knowledge of the hole distribution both in space and time. To obtain analytical expressions for the weight functions and the charging currents, the authors restricted attention to quasi-static changes of electron and hole distributions in the

quasi-neutral base layer [Klo87] They emphasized small-signal analysis in their illustration but they indicate that extensions to large-signal transients are possible The resulting TLCC formulation simulates the phase of $y_{21}(\omega)$ better than the Gummel-Poon formulation does The standard used for comparison is the exact solution for uniform base doping and low-level injection [Klo87]. In addition they connect the small-signal TLCC model to the time delay approach proposed by Te Winkel [Tow73] The TLCC relation is then assessed and extended to high injection from the numerical data obtained by MEDUSA [Eng83]

6 2 3 Negative-Capacitance (NC) and Inductive (L) Equivalent Circuits

I sketch here the approach developed to include the effect of propagation delay in quasi-neutral layers The NC and L equivalent circuits has been derived in details in Chapter Five The resulting NC and L models are applicable to all doping profiles to all levels of injection and to saturated and forward-active operation To simplify the discussion, I make the following assumptions (a) low injection (b) negligible base recombination (c) forward active operation and (d) exponential doping profile in the base layer, that is, $N_A(x)/N_A(0) = \exp(-\eta x/X_B)$ where x=0 defines the emitter edge of the quasi neutral base (Many transistors have an approximately exponential doping profile over the greater part of the base layer, although the profile in some transistors leads to a retarding built-in electric field near the base edge of the emitter-base space charge layer) The derivation begins with the electron continuity equation Laplace transformation yields a second-order ordinary differential equation in position x in

which the complex frequency variable s is treated as a parameter The solution of this ordinary differential equation for the transformed excess electron density involves two arbitrary constants To remove these constants I write equations for the excess electron density and electron current density at the two edges of the quasi-neutral base layer yielding four equations from which the two constants can be eliminated The result is a 2X2 matrix equation for the current density at each edge as a function of the excess electron density or the quasi-Fermi-potential separation at the edges This equation resembles those appearing in two port network theory As is desirable for models for circuit simulation no dependence on space variables appears But the distributed (partial-differential-equation) nature of the quasi neutral base remains imbedded in the transcendental rather than the algebraic functions of s that constitute the elements of the 2X2 matrix The elements appear in Eq (5 5)-(5 8) of Chapter Five To obtain lumped equivalent-circuit models, I truncate expansions of these transcendental functions Several such approximations are possible One approximation, (5 16), yields the quasi static model A second approximation, (5 19), leads to a quasi-static equivalent circuit including a negative capacitor (NC), shown in Fig 5 3(a) An alternate truncation (5 22), replaces the negative capacitor by an inductor, yielding the inductive or L equivalent circuit in Fig 5 3(b) The L equivalent circuit provides greater accuracy for a given number of circuit elements Note that the nodal variable is n_B in the equivalent circuits in Fig 5 3 rather than voltage as in the conventional equivalent circuit The equivalent circuits in Fig 5 3

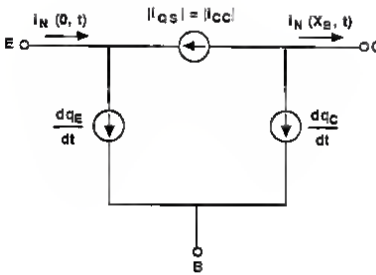


Fig 6 2 The charge-controlled representation of the NC model and the L model For detailed derivation of this equivalent circuit, see Appendix C This is the unifying form for all intrinsic BJT models

serve as a stepping stone to help derive the conventional large and small signal L and NC equivalent circuits which are shown in Fig 5 4 and Fig 5 6, respectively Furthermore they facilitate the comparison of the different approximations The method described in Chapter Five generalizes recent work by Chen Lindholm and Jung [Che85] for small signal modeling

6 3 Comparison of FCR, TICC and NC Models

6 3 1 Large-Signal Comparison

I focus on the quasi-neutral base layer and neglect recombination for simplicity Base recombination can be easily included I also assume (1) no heavy doping effects, (2) low injection, and (3) forward-active operation

To begin the comparison I transform the NC model (Fig 6 3) to a controlled-charge representation (Fig 6 2) by the method explained in Appendix C From Fig. 6.2,

$$i_N(0,t) = -\left|i_{QS}(t)\right| + dq_E/dt \tag{6 1}$$

$$i_N(X_B,t) = \left|i_{QS}(t)\right| + dq_C/dt \tag{6 2}$$

and

$$q_{QS} = q_E + q_C \ , \tag{6 3}$$

in which i_{QS}, q_C and q_E are defined in (5 29), (C 2) and (C 4) for an exponentially doped base layer Uniform doping corresponds to the special case $\eta = 0$

Equations (6 1) through (6 3) together with (5 29) (C 2) and (C 4) also describe the TICC model The TICC model was not illustrated for an exponentially doped base layer [Kle87] Thus I provide a derivation in Appendix D The result is (D 7) (D 8), and (D 9) which are equivalent to (C 2), (C 4) and (5 29) This proves that the NC and TICC models are equivalent for an exponentially doped base layer

Equations (6 1) through (6 3) together with (5 29) (C 2) and (C 4) describe the PCB model only for the special case of a uniformly doped base layer As was suggested in Fossum and Veeraraghvan [Foe86] the PCB model applies only approximately for an exponentially doped base and the inaccuracy involved for such a base was not given In Appendix D, I derive the PCB model for an exponentially doped base to estimate the inaccuracy For such a base, in contrast to the conduction currents for the NC and TICC models (5 29 and D 3), the conduction current for the PCB model (D 14) requires knowledge of the total time varying excess electron density $n(x,t)$ in the quasi-neutral base In Appendix D, therefore, I approximate $n(x,t)$ by its quasi-static counterpart $n_{QS}(x,t)$ The results are shown in Table 6 1 and Fig 6 3 Table 6 1 shows i_{QS} and i_C normalized by i_{QS} for the PCB TICC and NC models In constructing Table 6 1, I assume an exponentially doped base layer in low injection To provide a standard for comparison, I calculate the exact $i_N(X_B,t)$ which is then separated into a steady-state portion and an a time dependent portion via Taylor expansion Thus

$$i_N(X_B,s) = i[i_N(X_B,s=0) + s\left.\frac{di_N(X_B,s)}{ds}\right|_{s=0} + \quad] \qquad (6\ 4)$$

	$q_{QS}/i_{QS} = \tau_{tr}$	$q_C/i_{QS} = \tau_{21}$	τ_{21}(exact) Eq-(6 5)
TIGG	$\frac{X_E^2}{D_n}\left(\frac{\eta}{\eta^2}\frac{1+e^{\eta}}{}\right)$	$\frac{X_E^2}{D_n}\left(\frac{1+e^{-\eta}}{\eta(1-e^{-\eta})} - \frac{2}{\eta^2}\right)$	$\frac{X_B^2}{D_n}\left(\frac{1+e^{\eta}}{\eta(1-e^{\eta})} - \frac{2}{\eta^2}\right)$
NC			
PCB		$\frac{X_B^2}{D_n}\left[\frac{1}{2\eta}-\frac{1}{\eta^2}+\frac{1}{\eta}\left(1-e^{\eta}\right)\right]$	

Table 6-1 q_{QS} and q_C normalized by i_{QS} for the TIGG NC and PCB models for an exponentially doped base, and the τ_{21}(exact) Note that $q_E = q_{QS} - q_C$ and is not given here The relation for τ_{21} from the TIGG and NC models is more accurate than that for the PCB model except when the field factor η is zero (uniformly doped base) then $\tau_{21} = X_B^2/6D_n$ by the L'Hopital rule

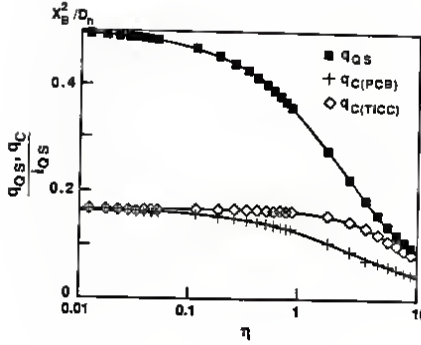


Fig. 6.3 Normalized q_{QS} and q_C of the TICC (or NC) and PCB models for an exponentially doped base layer with different values of η

Extracting $-i_N(X_B, s=0)$ which is $|i_{QS}|$, from (6.4) yields

$$-i_N(X_B, s) = |i_{QS}| \left(1 + s\, r_{21(exact)} + \quad \right) \quad \ldots$$

where $r_{21(exact)} = \dfrac{\left.\dfrac{di_N(X_B, s)}{ds}\right|_{s=0}}{|i_{QS}|}$ (6.5)

The parameter $r_{21(exact)}$ is used as a standard for comparison. From Table 6.1, I see that the TICC and NC models predict the correct $r_{21} = q_C/i_{QS}$ the PCB model underestimates q_C by 50% for $\eta = 8$ (Fig. 6.3).

For a uniformly doped base layer, $FE(x) = 1$, x/X_B and $FC(x) = x/X_B$, obtained by putting $\eta = 0$ in (D.5) and (D.6) and by the L'Hopital rule. Thus (D.3) and (D.4) are equivalent to (D.10) and (D.11), which shows that the TICC model is equivalent to the PCB model. As a result, the NC, PCB and TICC models are equivalent for a uniformly doped base layer.

In summary, the NC and TICC models are equivalent for an exponentially doped base layer having an arbitrary field factor η and are described by (6.1)-(6.3) together with (5.29), (D.2) and (D.4). If $\eta = 0$, which is the special case corresponding to a uniformly doped base layer, these relations then also describe the PCB model.

6.3.2 Small Signal Comparison

Small-signal analysis is a special case, directly derivable from large-signal analysis. Thus the large signal comparison explained above applies as well for the small-signal versions of the PCB, TICC and NC models.

6.4 L Model

The L model differs from the other three models and is not compared in Section 6.3. To compare the L model with the NC model, I first explore small-signal excitation as preparation for the large signal comparison. Because I have already demonstrated the relations among the TICC and PCB models, treating the L and the NC models provides a comprehensive comparison.

6.4.1 Small-Signal Comparison

I compare small-signal transconductances, y_{21}, predicted by the NC and L models. The small-signal NC and L models are derived in Chapter Five. The standard used for the comparison is the exact y_{21} for an exponentially doped base layer [Soh70].

From Fig. 6.4(a), I see that the L model simulates the magnitude of y_{21}, $|y_{21}|$, better than the NC model does. Figure 6.4(b) shows that the NC and L models predict the same phase of y_{21}.

Next, I discuss why the small-signal (ss) L model is the most accurate of the four models. The exact $y_{21}(s)$ has an infinite number of left-half-plane (LHP) poles and no zeros [Lyn67]. The $y_{21}(s)$ of the NC model (and of the TICC and PCB models) has a right half plane (RHP) zero and no poles:

$$y_{21(NC)} = i_c(s)/v_{be} = (1/R_{\pi}^{ss}) - s\,C^{ss} \quad , \tag{6.5}$$

as Fig. 6.4(a) implies. In contrast, the level of approximation in the L model leads instead to a single LHP pole:

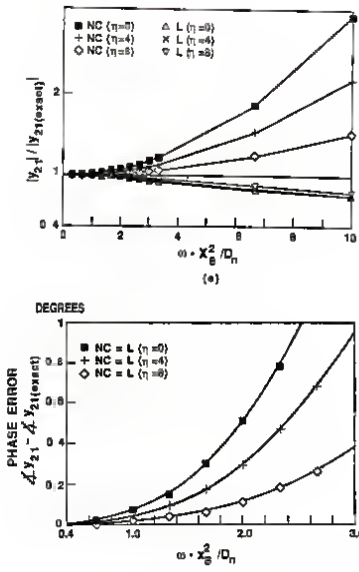


Fig. 6.4 Comparison of small signal transconductance predicted by the TICC and NC models. The standard for comparison is the exact solution for an exponentially doped base and low level injection. (a) Ratio of $|y_{21}|$ predicted to the exact $|y_{21}|$ (b) phase difference between the predicted and the exact value

$$y_{21(L)} = L/(R_{qs}^{se} + s\,L^{se}) \qquad (6.7)$$

Although the RHP zero predicts the same phase as the LHP pole does, it predicts that $|y_{21}|$ rises with frequency ($d|y_{21}(\omega)|/d\omega > 0$), which is incorrect. This is shown in Figure 6.5 in which for the exact $|y_{21}|$, $d|y_{21}(\omega)|/d\omega < 0$. This contrasts with the predictions of the NC model but agrees with the predictions of the L model. That $d|y_{21}(\omega)|/d\omega > 0$ in the NC model and its equivalents may limit the range of model applicability is further discussed in Section 6.5.

The L model gives the correct input capacitance and provides an accurate simulation of the exact input admittance, y_{11}, for $\omega < 1/\tau_{tr}$, as is shown in Chen and coworkers [Che86]. The NC model has similar accuracy for y_{11} but is not shown in Chen and coworkers [Che86].

6.4.2 Large-Signal Comparison

Large-signal analysis requires the use of user-defined controlled sources, which is the subject of Chapter Seven, to implement the nonlinear current components in a circuit simulator. As will be shown in Section 7.6, I apply a unit step voltage as the input to a transistor. The output current predicted by the NC model shows a negative impulse at the time when the input is applied (Fig. 7.2). The discrepancy is related to the incorrect $d|y_{21}(\omega)|/d\omega > 0$ predicted by the NC model. The simulation uses the L and NC models requires comparable CPU time (Section 5.5.2).

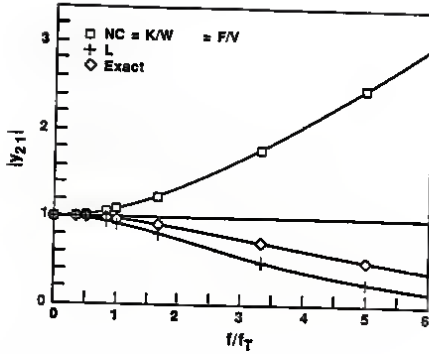


Fig 6.5 Comparison of small signal y_{21} for $\eta = 0$ predicted by the TICC and NC models. The standard for comparison is the exact solution for an exponentially doped base and low level injection. The NC model predicts that $|y_{21}|$ rises with frequency ($d|y_{21}(\omega)|/d\omega > 0$) which is incorrect. In contrast, for the exact $|y_{21}|$, $d|y_{21}(\omega)|/d\omega < 0$

6.5 PCE, TICC and NC Models. Possible Fundamental Limitations

I discuss the physical significance of the rise of $|y_{21}|$ predicted by the NC TICC and PCE models To simplify the quantitative discussion, I shall assume a uniformly doped base layer The result is, however, applicable to all doping profiles

6.5.1 Fundamental Limitations of PCE, TICC and NC Models

Paulos and Antoniadis [Pau83] examined the partitioned charge models in MOSFET and pointed out that such models which have transcapacitances, can be thermodynamically unsound [Pau83] According to them circuits with transcapacitors produce net energy and the power generation eventually becomes larger than the net quiescent power dissipation at very high frequencies In terms of device characteristics this could imply an upper frequency limit for the validity of the model and possible problems in transient circuit simulations. The partition of the base charge which is employed by the PCE TICC and the equivalent NC models, also leads to transcapacitors and hence to similar limitations The L model has no such limitation.

6.5.2 Quantitative Explanations

Paulos' explanation is qualitative I shall give a quantitative explanation for the undesired rise of $|y_{21}|$ predicted by the NC model, in the sense that I define its applicable frequency range The quantitative discussion makes use of the derivation in Chapter Five

Chapter Five shows three approximations of Z csch(Z) which are (5 16) (5 18) and (5 22) The Taylor expansion (5 18) which yields the NC model, converges provided that

$$|Z| < \pi \qquad (6\ 8)$$

Thus the interval of convergence of the NC model for a uniformly-doped base layer is

$$\omega < \pi\ D_n/X_B^2 \qquad (6\ 9)$$

This explains the rise of $|y_{21}|$ predicted by the NC model once the frequency approaches the limit imposed by the interval of convergence To illustrate a consequence of this limitation, I apply a small-signal step voltage as the input to an intrinsic transistor (Fig 6 6) In contrast to the exact solution, the output current predicted by the NC model shows a negative impulse at the time when the input is applied (Fig 6 4) The discrepancy is related to the incorrect $d|y_{21}(\omega)|/d\omega > 0$ predicted by the NC model

6 5 3 Solutions to the Limitation of PCE, TICC and NC Models

In contrast to (5 19) the Taylor expansion (5 22), which leads to the L model, converges for all Z As a result, the L model produces more accurate results both in frequency-domain (Fig 6 4 and 6 5) and time-domain (Fig 6 6) simulations I therefore recommend using the L model to circumvent the fundamental limitation of the NC model (and the

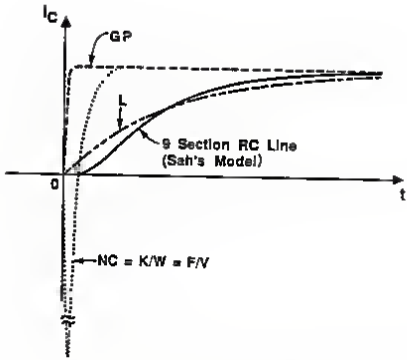


Fig 6 6 The collector current response to a small signal step ideal input voltage derived via the L model and via the NC model compared with the nine section rc transmission line [Sah70] as the exact solution for low level injection This is a severe test for all models Note that the response predicted by Gummel-Poon model reproduces the input excitation Note that K/W and F/V are Klose and Wieder [Kle87] and Fossum and Veeraraghvan [Fos86] respectively

PCB and TICC models). Moreover, the L model has a wider applicable frequency range, which is determined by requiring that the second term greatly exceed the third term of the Taylor expansion shown in (5.19) and (5.22). Thus, for (5.19), the range is

$$|Z^2/6| \gg |7\,Z^4/360|, \tag{6.10}$$

and for (5.22), the range is

$$|Z^2/6| \gg |Z^4/120|. \tag{6.11}$$

For a uniformly doped base layer (5.19) is valid for

$$\omega \ll 60\,D_n/(7\,X_B^2), \tag{6.12}$$

and (5.22) is valid for

$$\omega \ll 20\,D_n/X_B^2 \tag{6.13}$$

Thus, the L model has about a three times larger applicable frequency range than the NC model has for a uniformly doped base layer.

6.6 Summary and Discussion

I compare various non-quasi-static (NQS) improvements of the Gummel-Poon integral charge control model for both small-signal and large-signal excitations. The comparison includes the partitioned

charge-based (FCB) model by Fossum and Veeraraghavan, the transient integral charge control model (TICC) by Klose and Wieder, and negative capacitance (NC) and inductive (L) NQS equivalent circuits derived by truncation in the complex-frequency domain. For forward-active operation and an exponential doping profile in the base layer, the NC model is equivalent to the TICC model. For a uniform doping profile, the NC, TICC and FCB models are equivalent. These three models, however, predict that the magnitude of the transconductance rises with frequency for high frequencies, a trend that, in principle, is incorrect. Truncation in the s domain gives insight concerning this trend and provides a better model: the inductive model or L model.

The method of truncation emphasizes solution of differential equations which the presence of quasi-neutral layers makes possible. Thus, the method of truncation is well suited for bipolar transistors; it is inappropriate for large-signal ($v > kT/q$) characterization of MOS transistors and other devices in which the space charge-limited flow prevails. The model of Cummel-Poon [Gum70b] as well as the FCB and TICC models, derives from integrating differential equations, as do all charge control models. The great utility of charge control models is their applicability to all devices, including MOS transistors. For large signal modeling of bipolar devices, the truncation method yields two forms of non-quasi-static models: one that includes negative capacitors and another that includes instead inductors. The negative capacitor versions are equivalent to the TICC and FCB model under certain restrictions. The inductive (L) versions are more accurate, yet no more complex.

CHAPTER SEVEN
COMPUTER IMPLEMENTATION OF THE LARGE SIGNAL L AND NC MODELS USING USER DEFINED CONTROL SOURCES

7 1 Introduction

The method of truncation in the complex frequency domain results in the NC and L models, which account for non-quasi-static (NQS) effects for both small and large signal excitation [Che88] Although these models work directly in SPICE for small-signal applications, their use in SPICE for large-signal dynamic applications requires either rewriting the code or using the user defined control source (UDCS) approach [Vee86, Yum87] In this chapter, I explore the UDCS approach in SLICE, which is a custom version of SPICE from Harris Semiconductor Corporation The UDCS is a convenient way to simulate nonlinear (transient) models in SLICE [Sli86]

The purposes of this chapter are as follows

(1) to discuss numerical aspects of using the large-signal L and NC models in a circuit simulation program

(2) to provide detailed documentation for future use of the L and NC models in circuit simulation programs and

(3) to employ the large-signal L and NC models to help measure the importance of non-quasi-static models in digital applications

I start with a brief review of related numerical techniques In Section 7 3, I discuss the large signal NC model implementation giving special attention to the current source $dq_{C(NC)}/dt$ which accounts for

the non-quasi static effects. In Section 7.4, I discuss the implementation of the current source $dq_{C(L)}/dt$ in the L model. I then apply the large-signal L and NC models to simulate a single-transistor inverter.

In addition to Vasireghvan's work [Vas86] on the UDCS methodology, Yuan and Eisenstadt [Yua87] described the incorporation of the Gummel-Poon bipolar transistor model into UDCS form for SLICE. We shall refer to their works wherever possible to avoid repetition.

7.2 Numerical Integration and Nonlinear Solution Algorithms

To provide a road map for simulating nonlinear transient circuits, I briefly describe two distinct aspects of the simulation: the nonlinear solution algorithm and the numerical integration algorithm. I also discuss a minor error in Vasireghvan and coworkers [Vas86].

Simulating a nonlinear circuit requires a nonlinear solution algorithm to determine the solution of the system of nonlinear equations. This algorithm is an iterative sequence of linearized-equations solutions. Simulating the transient nonlinear system also requires a numerical integration algorithm to reduce the system of differential circuit equations, at each time point, to a system of algebraic equations. For details, see Chua and Lin [Chu75].

7.2.1 Nonlinear Solution Algorithm

SLICE/SPICE uses branch linearization for the circuit equations, which allows formulating each individual branch rather than the overall system of nonlinear equations. I illustrate the method for branch

linearization by the Newton-Raphson algorithm which linearizes nonlinear current equations $i=f(v)$ by the Taylor series expansion

$$i(v^{m+1}) = i(v^m) + \left.\frac{\partial i}{\partial v}\right|_{v^m}(v^{m+1} - v^m)$$

$$= \left.\frac{\partial i}{\partial v}\right|_{v^m} v^{m+1} + [i(v^m) - \left.\frac{\partial i}{\partial v}\right|_{v^m} v^m] \qquad (7.1)$$

If the branch relation is current defined and depends only upon its own branch voltage, then the linearization procedure in (7.1) is equivalent to replacing the nonlinear branch, at each iteration, by a linear conductance

$$\partial i/\partial v\Big|_{v^m}$$

in parallel with an independent current source

$$i(v^m) - \partial i/\partial v\Big|_{v^m} v^m$$

7.2.2 Numerical Integration Algorithm

Determining the transient response over some specific time (0,T) requires a numerical integration method. Consider, as an example, the nonlinear differential equation,

$$i = d[q[v(t)]]/dt \qquad (7.2)$$

The first step is to reduce (7.2) to an algebraic equation using a numerical integration algorithm at each time point. For example,

applying the implicit trapezoidal algorithm to (7 2) at time point t = t_{n+1} yields:

$$i(t_{n+1}) = \frac{2\,q(t_{n+1})}{\Delta t} - [\frac{2\,q(t_n)}{\Delta t} + i(t_n)] \qquad (7\ 3)$$

where $q(t_n)$ and $i(t_n)$ are either the initial conditions at the first time point or solutions from the last time point. The choice of a numerical integration method involves a trade-off of the local truncation error and stability. The implicit trapezoidal and the Gear-2 algorithms are used in SPICE2 [Nag75].

In general Eq. (7 3) is a nonlinear equation. The linearization of Eq. (7 3) is done similarly to that of Eq. (7 1). The result after linearization is given in terms of "transconductance" and "independent current sources" as shown in (7-4) and (7-5), respectively.

$$G = \frac{2}{\Delta t}\,\frac{dq[v^{m-1}(t_{n+1})]}{dv} \qquad (7\ 4)$$

$$i = -\frac{2}{\Delta t}\,\frac{dq[v^{m-1}(t_{n+1})]}{dv}\,v^{m-1}(t_{n+1}) + \frac{2}{\Delta t}\,q[v^{m-1}(t_{n+1})]$$

$$[\frac{2\,q(t_n)}{\Delta t} + i(t_n)] \qquad (7\ 5)$$

Although some charge-based current sources in the MC and L models are not functions of their own model voltages, I still use the "independent current source" and "transconductance" to illustrate the concept of linearization.

In Veeraraghvan et al [Vee86], the expression for the "independent current sources" omitted the first term

$$-\frac{2}{\Delta t}\frac{dq[v^{n-1}(t_{n+1})]}{dv}v^{n-1}(t_{n+1}) \qquad (7.6)$$

Note, however, that this missing term (7.6), will not effect the UDCS simulation because a UDCS requires the user to supply the nonlinear algebraic equation, (7.3) and its voltage derivative instead of (7.4) and (7.5). Thus the UDCS takes care of the linearization internally.

7.3 Implementation of the NC Model

The intrinsic portion of the large-signal NC model shown as a charge based representation in Fig. 7.1, has three current sources: the nonlinear conduction current source, i_{QS}, and the nonlinear transient current sources dq_C/dt and dq_E/dt. I shall implement only i_{QS} and dq_C/dt, since dq_C/dt and dq_E/dt are similar. The implementation of the extrinsic portion of the NC model will not be shown here; it has been presented by Yuen and Eisenstadt [Yue87]. I shall also compare two numerical integration algorithms: the implicit trapezoidal and Gear-2 methods. The accuracy of the UDCS simulation is assessed by comparing with an analytical solution for a uniformly doped base.

7.3.1 Implementation of i_{QS}

Implementation of i_{QS}, which is the Gummel formulation for the conduction current (Eq. 5.29), is straightforward. The iterative procedure was described in Section 7.2.1 and is not repeated here. I

emphasize that the UDCS requires the user to provide (7 3) and its voltage derivative instead of the corresponding linearized currents as shown in (7 4) and (7 5)

To assess the accuracy of the simulation, I apply a input voltage $v_{gg}(t)$, which is a ramp starting at $t = 0$ and leveling off at $v = V_F - 0.8$ V at $t = TR$ with a slope a $(= V_F/TR)$ I choose the value of TR to be one tenth of the base transit time τ_{cr} This input voltage waveform is shown in Fig 7 1 To provide a standard for comparison, I substitute $v_{gg}(t)=a\,t$ into (5 29) to obtain an analytical solution for a uniformly doped base The accuracy of the simulated i_{QS} is excellent (Table 7 1) The accuracy of i_{QS} drops however as a lower time resolution for sweeping time points is used Since simulating i_{QS} does not require numerical integration, this degradation is caused by other time dependent current sources in the NC model for example, $dq_{C(NC)}/dt$ and $dq_{E(NC)}/dt$

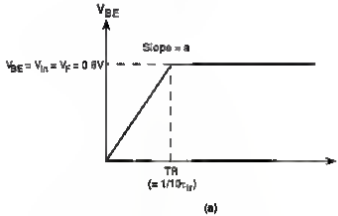
7 3 2 Implementation of dq_C/dt Using the Implicit Trapezoidal Method, Oscillation

Employing the implicit trapezoidal algorithm as the numerical integration algorithm to simulate $dq_{C(NC)}/dt$ (7 7), I get

$$i = \frac{dq_{C(NC)}}{dt} = \frac{d}{dt}\left[\tau_D \frac{ng(t)}{\bar{R}_{QS}}\right] \qquad (7\ 7)$$

Time	$i_{QS}(t)$				
τ_{tr}	TR/800	TR/400	TR/200	TR/100	Analytical
0.000	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.000E+00
0.004	1.647E-18	1.647E-18	1.648E-18	1.651E-18	1.647E-18
0.008	7.328E-18	7.328E-18	7.329E-18	7.340E-18	7.328E-18
0.012	2.692E-17	2.692E-17	2.692E-17	2.696E-17	2.692E-17
0.016	9.447E-17	9.447E-17	9.448E-17	9.461E-17	9.447E-17
0.02	3.274E-16	3.274E-16	3.275E-16	3.279E-16	3.274E-16
0.024	1.131E-15	1.131E-15	1.131E-15	1.132E-15	1.131E-15
0.028	3.901E-15	3.901E-15	3.901E-15	3.907E-15	3.901E-15
0.032	1.345E-14	1.345E-14	1.346E-14	1.347E-14	1.345E-14
0.036	4.639E-14	4.639E-14	4.640E-14	4.646E-14	4.639E-14
0.04	1.600E-13	1.600E-13	1.600E-13	1.602E-13	1.600E-13
0.044	5.517E-13	5.517E-13	5.516E-13	5.525E-13	5.517E-13
0.048	1.902E-12	1.902E-12	1.903E-12	1.905E-12	1.902E-12
0.052	6.560E-12	6.560E-12	6.562E-12	6.570E-12	6.560E-12
0.056	2.262E-11	2.262E-11	2.263E-11	2.266E-11	2.262E-11
0.06	7.801E-11	7.801E-11	7.803E-11	7.813E-11	7.801E-11
0.064	2.690E-10	2.690E-10	2.691E-10	2.694E-10	2.690E-10
0.068	9.277E-10	9.277E-10	9.278E-10	9.291E-10	9.277E-10
0.072	3.199E-09	3.199E-09	3.200E-09	3.204E-09	3.199E-09
0.076	1.103E-08	1.103E-08	1.103E-08	1.105E-08	1.103E-08
0.08	3.804E-08	3.804E-08	3.805E-08	3.810E-08	3.804E-08
0.084	1.312E-07	1.312E-07	1.312E-07	1.314E-07	1.312E-07
0.088	4.524E-07	4.524E-07	4.524E-07	4.530E-07	4.524E-07
0.092	1.560E-06	1.560E-06	1.560E-06	1.562E-06	1.560E-06
0.096	5.379E-06	5.379E-06	5.380E-06	5.387E-06	5.379E-06
0.1	1.855E-05	1.855E-05	1.855E-05	1.855E-05	N/A
0.12	1.855E-05	1.855E-05	1.855E-05	1.855E-05	N/A
0.14	1.855E-05	1.855E-05	1.855E-05	1.855E-05	N/A
0.16	1.855E-05	1.855E-05	1.855E-05	1.855E-05	N/A

Table 7.1 The simulation of i_{QS} by four different sweeping time resolutions TR/800, TR/400, TR/200 and TR/100, where TR (= τ_{tr}/10) is the rise time for the voltage ramp input. The accuracy of i_{QS} drops for a lower sweeping-time resolution. Since simulating i_{QS} does not require numerical integration, this degradation is caused by other time dependent current sources in the NC model, for example, $dq_{C(NC)}/dt$ and $dq_{E(NC)}/dt$. The time scale is normalized to τ_{tr}.



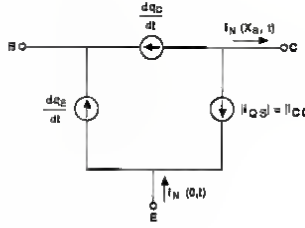


Fig 7.1 The input waveform of the input voltage and the charge based representation of the NC and L models. For detailed derivation of this equivalent circuit, see Appendix C. All elements depend on $v_{eg}(t)$.
(a) Input waveform (b) The equivalent circuit.

where $n_E = \frac{n_i^2}{N_A(0)}\{\exp(\frac{v_{BE}}{kT/q}) - 1\}$ R_{QS} and r_D are given in

(5 18) and (5 21) respectively

To provide a standard for comparison I substitute $v_{BE}(t) = e_1 t$ into (7 7) to obtain an analytical solution (7 8), for $dq_{C(NG)}/dt$ for a uniformly doped base

$$I(t) = \frac{dq_{C(NG)}}{dt} = \frac{A\, q\, D_n\, n_i^2}{N_A(0)\, X_B}\left[e^{\frac{e_1 t}{kT/q}} - \frac{e_1}{kT/q}\right] \qquad (7\ 8)$$

Note that analytical solutions for t > TR where the input voltage levels off are not available because superposition is not valid for a nonlinear system

The simulated and analytical results of $dq_{C(NG)}/dt$ are given in Table 7 2 From Table 7 2 I observe the following

(1) Good accuracy of simulated $dq_{C(NG)}/dt$ is obtained for t < TR

(2) The simulated $dq_{C(NG)}/dt$ oscillates at the same value with different sign when t > TR Here TR is the time that v_{BE} levels off at V_F = 0 8 V (Fig 7 1) The magnitude of the oscillation is about 1000 times of $I_{QS}(t = TR)$ This oscillation is incorrect because, for t > TR, $dq_{C(NG)}/dt$ should be zero, since $q_{C(NG)}(t)$ = constant for t > TR when the voltage flattens out at 0 8 V

To explain why the implicit trapezoidal method (7 9), predicts the oscillation at t > TR consider

Time / τ_{tr}	$da_{C(NC)}/dt$		
	implicit trapezoidal	Gear 2	Analytical
0 000	0 000E+00	0 000E+00	6 940E-17
0 004	2 865E-16	2 392E-16	2 393E-16
0 008	8 724E 16	8 249E-16	8 253E-16
0 012	2 893E 15	2 845E 15	2 846E-15
0 016	9 860E-15	9 810E-15	9 814E-15
0 020	3 389E 14	3 383E-14	3 384E-14
0 024	1 167E-13	1 167E-13	1 167E 13
0 028	4 025E-13	4 023E-13	4 024E-13
0 032	1 388E-12	1 387E 12	1 388E 12
0 036	4 785E-12	4 784E-12	4 786E-12
0 040	1 650E 11	1 650E-11	1 650E-11
0 044	5 690E 11	5 688E 11	5 691E-11
0 048	1 962E-10	1 962E-10	1 962E-10
0 052	6 767E-10	6 764E-10	6 767E-10
0 056	2 333E-09	2 333E-09	2 334E-09
0 060	8 047E 09	8 044E 09	8 047E-09
0 064	2 775E 08	2 774E-08	2 775E-08
0 068	9 569E 08	9 565E-08	9 570E-08
0 072	3 300E 07	3 298E-07	3 300E-07
0 076	1 138E 06	1 137E-06	1 138E-06
0 080	3 924E 06	3 922E 06	3 924E 06
0 084	1 353E 05	1 353E-05	1 353E-05
0 088	4 666E 05	4 664E-05	4 666E-05
0 092	1 609E 04	1 608E 04	1 609E 04
0 096	5 548E-04	5 546E-04	5 549E-04
0 100	1 123E-02	7 874E-03	1 913E-03
0 1001	5 978E 02	3 611E-13	N/A
0 1002	-3 146E 02	3 411E-13	N/A
0 1004	2 517E-02	3 411E-13	
0 1005	-2 517E-02	3 411E-13	
0 1006	2 517E-02	3 411E-13	
0 1007	-2 517E 02	3 411E 13	
0 1009	2 517E 02	3 411E-13	
0 1010	-2 517E 02	3 411E-13	
0 1011	2 517E 02	3 411E-13	
0 1012	-2 517E 02	3 411E-13	
0 1014	2 517E 02	3 411E-13	

Table 7 2 The simulation results of $da_{C(NC)}/dt$ by the implicit trapezoidal and the Gear 2 method The time scale is normalized to τ_{tr} Note that after the voltage levels off at 0 8V (at $t = 1.0\ \tau_{tr}$) $da_{C(NC)}/dt$ as predicted by the implicit trapezoidal method oscillates

$$i(t_{n+1}) = \frac{2\,q(t_{n+1})}{\Delta t} \quad (\frac{2\,q(t_n)}{\Delta t} + i(t_n)) \tag{7 9}$$

At $t = t_{n+1} > TR$, $v_{BE}(t_n) = v_{BE}(t_{n+1})$; therefore the first two terms in (7 9) are equivalent. As a result, the first two terms in (7 9) cancel and (7 9) becomes

$$i(t_{n+1}) = -i(t_n) \tag{7 10}$$

At the next time point, $t=t_{n+2} > TR$, $q(t_{n+2})$ and $q(t_{n+1})$ cancel because $v_{BE}(t_{n+2}) = v_{BE}(t_{n+1})$. Thus

$$i(t_{n+2}) = -i(t_{n+1}) = i(t_n) \tag{7 11}$$

For $t > TR$, i changes at alternate time points at the same value with different sign as observed in Table 7.2.

In summary, the oscillation is caused by the sudden leveling-off of the input voltage at $t = TR$ as well as the quasi-static change of $q_{C(NC)}$ with v_{BE}. This quasi-static relation between $q_{C(NC)}$ and v_{BE} for the NC model is obvious in (7 7). I shall employ the Gear 2 method to simulate $dq_{C(NC)}/dt$ to further investigate the oscillation problem.

7 3 3 Implementation of $dq_{C(NC)}/dt$ Using the Gear 2 Method

I employ the Gear-2 method instead of the implicit trapezoidal method to reduce differential equations at each time point, to difference equations. The result of simulating $dq_{C(NC)}/dt$ using the

Gear-2 method is given in Table 7.2. Observe that $dq_{C(NC)}/dt$ is approximately zero for $t > TR$. This is correct because $q_{C(NC)}$ = constant for $t > TR$. The Gear-2 method, (7.12), correctly predicts a zero current when v_{BE} levels off.

$$i(t_{n+1}) = \frac{3\, q(t_{n+1})}{\Delta t} - \frac{2\, q(t_n)}{\Delta t} + \frac{q(t_{n-1})}{2\,\Delta t} \qquad (7.12)$$

Because $q(t_n) = q(t_{n+1}) = q(t_{n+2})$ for $t > TR$, $dq_{C(NC)}/dt$ predicted by the Gear-2 algorithm is zero for $t = TR$. For simplicity, I assume uniform time steps for the Gear-2 method, in the actual implementation, I use the variable-time-step Gear-2 method.

From Table 7.2, I see that the simulation results of $dq_{C(NC)}/dt$ by the implicit trapezoidal method and the Gear-2 method are accurate for $0.01\, \tau_{tr} < t < 0.95\, \tau_{tr}$. The implicit trapezoidal method has a slightly better accuracy. However, I recommend using the Gear-2 method for the simulation of the NC model since the implicit trapezoidal method predicts an undesirable oscillation.

7.4 Implementation of the L Model

A charge-based representation of the large-signal L model is given in Fig. 7.1. UDCS simulation of the L model is similar to that of the NC model except that $dq_{C(L)}/dt$, given in (5.40) and repeated in (7.13), involves an integration.

$$q_{C(L)} = \frac{e^{-t/\tau_D}}{R'_{QS}} \int_0^t e^{\xi/\tau_D}\, v_E(\xi)\, d\xi - i(0)\, \tau_D\, e^{-t/\tau_D} \qquad (7.13)$$

Although the simulation of $dq_{C(L)}/dt$ is not straightforward because of the integration I shall show that evaluating the integral does not require an additional subroutine Since the implicit trapezoidal method results in oscillation in simulating $dq_{C(NC)}/dt$, I shall first use the Gear-2 method and then the implicit trapezoidal method to simulate $dq_{C(L)}/dt$

7 4 1 Implementation of $dq_{C(L)}/dt$ Using the Gear-2 Method

The numerical integration algorithm divides the simulation time [0, T] into individual time points [0, t_1, t_2, t_3 ... t_n t_{n+1} ... T] I use the Gear-2 algorithm, (7 12), to reduce the differential equation, $i_L = dq_{C(L)}/dt$, to an algebraic equation at each time point Since the three terms in (7 12) are similar only the first term in (7 12) is used for demonstration I now illustrate that no additional subroutine is required to evaluate the integral by going through the first three time points 0 t_1 and t_2 I then extend the result to the time point t_{n+1} as a summary

The first three time points 0, t_1, t_2 At $t = t_1$, (7 13) becomes

$$q_{C(L)} = \frac{-t_1/\tau_D}{R'_{QS}} \int_0^{t_1} a^{t/\tau_D} n_E(t)\, dt - i(0)\, \tau_D\, a^{-t_1/\tau_D} \qquad (7\ 14)$$

Approximating the integral by the trapezoidal integration method as shown in (7 15),

$$\int_0^{t_1} a^{t/\tau_D} n_E(t)\, dt = \frac{\Delta t}{2} \left[a^{t_1/\tau_D} n_E(t_1) + n_E(0) \right] \qquad (7\ 15)$$

I obtain an algebraic equation for (7 14) This result is stored

At the second time point $t = t_2$ (7 13) becomes

$$q_{C(L)} = -\frac{e^{-t_2/\tau_D}}{R'_{QS}} \int_0^{t_2} e^{\zeta/\tau_D} n_E(\zeta)\, d\zeta - i(0)\, \tau_0\, e^{-t_2/\tau_D} \qquad (7\ 16)$$

I separate the integration limit $(0, t_2)$ into $(0, t_1)$ and (t_1, t_2) as shown in (7 17)

$$\int_0^{t_2} e^{\zeta/\tau_D} n_E(\zeta)\, d\zeta = \int_0^{t_1} e^{\zeta/\tau_D} n_E(\zeta)\, d\zeta + \int_{t_1}^{t_2} e^{\zeta/\tau_D} n_E(\zeta)\, d\zeta \qquad (7\ 17)$$

In (7 17), the first integral is equivalent to (7 15) which was stored previously The second term can be handled in a similar way (See Eq 7 15)

Generalization to $t = t_{n+1}$ As a summary, I now generalize the above illustration to $t = t_{n+1}$ At $t = t_{n+1}$, as in the procedure at t_1 and t_2, I separate the integration limit $(0, t_{n+1})$ into $(0, t_n)$ and (t_n, t_{n+1}) which yields (7 18)

$$q_{C(L)} = -\frac{e^{-\frac{t_{n+1}}{\tau_0}}}{R'_{QS}} \left[\int_0^{t_n} e^{\frac{\zeta}{\tau_0}} n_E(\zeta)\, d\zeta + \int_{t_n}^{t_{n+1}} e^{\frac{\zeta}{\tau_D}} n_E(\zeta)\, d\zeta\right] - i(0)\, \tau_0\, e^{-t_{n+1}/\tau_D} \qquad (7\ 18)$$

The first integral in the brackets of (7 18) is known by keeping track of the solutions from previous time points The second integral can be approximated by the trapezoidal integration method

$$\int_{t_n}^{t_{n+1}} e^{\zeta/r_D}\, n_E(\zeta)\, d\zeta = \frac{\Delta t}{2}\left[e^{\frac{t_{n+1}}{r_D}}\, n_E(t_{n+1}) + e^{\frac{t_n}{r_D}}\, n_E(t_n)\right] \qquad (7\ 19)$$

where $\Delta t = t_{n+1} - t_n$. The approximation in (7 19) is accurate because Δt is kept small to maintain an acceptable local truncation error of the numerical integration algorithm. Putting (7 19) into (7 18) yields (7 20)

$$q_{C(L)} = \frac{e^{-\frac{t_{n+1}}{r_D}}}{R'_{DS}}\left[\int_0^{t_n} e^{\frac{\zeta}{r_D}}\, n_E(\zeta)\, d\zeta + \frac{\Delta t}{2}\left[e^{\frac{t_{n+1}}{r_D}}\, n_E(t_{n+1}) + e^{\frac{t_n}{r_D}}\, n_E(t_n)\right]\right]$$

$$+ i(0)\, r_D \exp(-t_{n+1}/r_D) \qquad (7\ 20)$$

Equation (7 20) is used as input to the UDCS simulation with some minor modifications, discussed below. This completes the proof that no integration subroutine is required to evaluate the integral in $dq_{C(L)}/dt$.

Since I only evaluate the integral from $t = t_n$ to $t = t_{n+1}$ instead of evaluating the integral from $t = 0$ to $t = t_{n+1}$, comparable CPU time is expected for simulating $dq_{C(L)}/dt$ and $dq_{C(NC)}/dt$.

7 4 1 1 Overflow problem

Equation (7 20) overflows at $t > r_D$ because of the exponential terms $\exp(t_n/r_D)$ and $\exp(t_{n+1}/r_D)$. To solve the overflow problem, I rewrite (7 20) as (7 21) by moving the term, $\exp(-t_{n+1}/r_D)$, into the bracket to cancel the exponential terms.

$$q_{G(L)} = \frac{1}{R'_{QS}} \left[e^{-\frac{t_{n+1}}{\tau_D}} \int_0^{t_n} e^{\frac{\zeta}{\tau_D}} n_E(\zeta)\, d\zeta + \frac{\Delta t}{2} \left[n_E(t_{n+1}) + e^{-\frac{\Delta t}{\tau_D}} n_E(t_n) \right] \right]$$

$$+ i(0)\, \tau_D \exp(-t_{n+1}/\tau_D) \qquad (7\ 21)$$

Using (7 21) instead of (7 20) as the input for UDCS prevents numerical overflow

7 4 1 2 Program subtleties

In the UDCS Fortran program (Appendix B) I use $\exp(-\Delta t/\tau_D)$ instead of $\exp(-t_{n+1}/\tau_D)$, which is the term in front of the integral in (7 21) This is not an error for the following reason In the FORTRAN program, what is stored at the last time point is the term shown in (7 22) rather than (7 23) Note that X(LP+21) in (7 22) is the variable used in FORTRAN program to stored the integration result

$$X(LP+21) = e^{-\frac{t_n}{\tau_D}} \int_0^{t_n} e^{\frac{\zeta}{\tau_D}} n_E(\zeta)\, d\zeta \qquad (7\ 22)$$

$$X(LP+21) = \int_0^{t_n} e^{\frac{\zeta}{\tau_D}} n_E(\zeta)\, d\zeta \qquad (7\ 23)$$

Rewriting (7 21) in terms of (7 22) yields (7 24)

$$q_{C(L)} = \frac{1}{R'_{QS}} \left[e^{\frac{\Delta t}{\tau_D}} X(LP+21) + \frac{\Delta t}{2} \left[ng(t_{n+1}) + e^{\frac{-\Delta t}{\tau_D}} ng(t_n) \right] \right]$$

$$i(0)\ \tau_D \exp(-t_{n+1}/\tau_D) \qquad (7\ 24)$$

From (7.24), I see that i should use $\exp(-\Delta t/\tau_D)$ instead of $\exp(-t_{n+1}/\tau_D)$ in front of X(LP+21) in fact it is impossible to store (7.23) because it overflows at $t > \tau_D$ The UDCS simulation of $dq_{C(L)}/dt$ using the Gear 2 method is given in Table 7 3

7 4 2 Implementation of $dq_{C(L)}/dt$ Using the Implicit Trapezoidal Method

I discuss use of the implicit trapezoidal method to implement $dq_{C(L)}/dt$ in contrast to the implementation of $dq_{C(NC)}/dt$ in the NC model the implicit trapezoidal method predicts a correct $dq_{C(L)}/dt$ (Table 7.4) no oscillation is observed after $t > TR$ This is because for $t_n > TR$ even though $v_{BE}(t_n) = v_{BE}(t_{n+1})$, $q_{C(L)}(t_n)$ is not necessarily equal to $q_{C(L)}(t_{n+1})$ in general in other words $q_{C(L)}$ does not follow v_{BE} quasi statically which is further discussed in the following

In the L model L appears in series with R which in circuit terms would give $i = (1/L)v/s$ or $(1/L)\int v\,dt$ for the response over a short time to rapid excitation (for which the voltage drops mainly across L) In physical terms v corresponds (for large-signals) to n_0 [exp(e v/kT) - 1] (or v is the electrochemical potential for

Time	$dq_{C(L)}/dt$ the Gear 2 method				
τ_{tr}	TR/800	TR/600	TR/200	TR/100	Analytical
0.000	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.000E+00
0.004	1.639E-18	1.656E-18	1.628E-18	1.602E-18	1.639E-18
0.008	7.270E-18	7.263E-18	7.254E-18	7.146E-18	7.273E-18
0.012	2.667E-17	2.644E-17	2.655E-17	2.624E-17	2.684E-17
0.016	9.556E-17	9.346E-17	9.515E-17	9.208E-17	9.359E-17
0.02	3.242E-16	3.239E-16	3.227E-16	3.191E-16	3.243E-16
0.024	1.119E-15	1.118E-15	1.114E-15	1.102E-15	1.120E-15
0.028	3.862E-15	3.858E-15	3.844E-15	3.801E-15	3.863E-15
0.032	1.332E-14	1.551E-14	1.326E-14	1.311E-14	1.332E-14
0.036	4.595E-14	4.588E-14	4.572E-14	4.521E-14	4.595E-14
0.04	1.584E-13	1.582E-13	1.577E-13	1.559E-13	1.584E-13
0.044	5.462E-13	5.456E-13	5.457E-13	5.376E-13	5.464E-13
0.048	1.885E-12	1.882E-12	1.875E-12	1.854E-12	1.884E-12
0.052	6.495E-12	6.488E-12	6.465E-12	6.395E-12	6.497E-12
0.056	2.240E-11	2.237E-11	2.230E-11	2.205E-11	2.240E-11
0.06	7.723E-11	7.716E-11	7.688E-11	7.602E-11	7.726E-11
0.064	2.665E-10	2.661E-10	2.651E-10	2.621E-10	2.664E-10
0.068	9.184E-10	9.175E-10	9.143E-10	9.040E-10	9.188E-10
0.072	3.167E-09	3.164E-09	3.153E-09	3.117E-09	3.168E-09
0.076	1.092E-08	1.091E-08	1.087E-08	1.075E-08	1.093E-08
0.08	3.766E-08	3.762E-08	3.749E-08	3.707E-08	3.767E-08
0.084	1.299E-07	1.297E-07	1.293E-07	1.278E-07	1.298E-07
0.088	4.479E-07	4.474E-07	4.458E-07	4.408E-07	4.480E-07
0.092	1.544E-06	1.543E-06	1.537E-06	1.520E-06	1.545E-06
0.096	5.326E-06	5.320E-06	5.301E-06	5.242E-06	5.327E-06
0.1	1.812E-05	1.788E-05	1.741E-05	1.655E-05	N/A
0.12	1.731E-05	1.732E-05	1.733E-05	1.736E-05	N/A
0.14	1.630E-05	1.651E-05	1.652E-05	1.655E-05	
0.16	1.555E-05	1.536E-05	1.537E-05	1.540E-05	
0.18	1.446E-05	1.447E-05	1.448E-05	1.450E-05	
0.2	1.562E-05	1.562E-05	1.564E-05	1.566E-05	

Table 7.3. The simulation of $dq_{C(L)}/dt$ using the Gear-2 method by four different sweeping time resolutions: TR/800, TR/400, TR/200 and TR/100, where TR ($= \tau_{tr}/10$) is the rise time for the voltage ramp input. The time scale is normalized to τ_{tr}

Time	$dq_{C(L)}/dt$, the implicit trapezoidal method				
τ_{tr}	TR/800	TR/400	TR/200	TR/100	Analytical
0 000	0 000E+00	0 000E+00	0 000E+00	0 000E+00	0 000E+00
0 004	1 639E-18	1 640E-18	1 640E-18	1 643E-18	1 643E-18
0 008	7 273E-18	7 274E-18	7 275E-18	7 287E-18	7 273E-18
0 012	2 668E-17	2 668E-17	2 669E-17	2 673E-17	2 668E-17
0 016	9 359E-17	9 359E-17	9 361E-17	9 375E-17	9 359E-17
0 02	3 243E-16	3 243E-16	3 244E-16	3 248E-16	3 243E 16
0 024	1 120E-15	1 120E-15	1 120E-15	1 122E-15	1 120E 15
0 028	3 863E-15	3 863E-15	3 864E-15	3 870E-15	3 863E-15
0 032	1 332E-14	1 332E-14	1 333E-14	1 335E-14	1 332E 14
0 036	4 595E-14	4 595E-14	4 596E-14	4 602E-14	4 595E-14
0 04	1 584E-13	1 584E-13	1 585E-13	1 587E-13	1 584E-13
0 044	5 464E-13	5 464E-13	5 465E-13	5 473E-13	5 464E-13
0 048	1 884E-12	1 884E-12	1 885E-12	1 887E-12	1 884E-12
0 052	6 497E-12	6 497E-12	6 499E-12	6 505E-12	6 497E-12
0 056	2 241E-11	2 241E-11	2 241E 11	2 244E 11	2 240E-11
0 06	7 726E-11	7 726E-11	7 728E-11	7 739E-11	7 726E-11
0 064	2 664E-10	2 664E-10	2 665E-10	2 669E-10	2 664E-10
0 068	9 188E-10	9 188E-10	9 190E-10	9 203E-10	9 188E-10
0 072	3 168E-09	3 168E-09	3 169E-09	3 174E-09	3 168E-09
0 076	1 093E-08	1 093E-08	1 093E-08	1 094E-08	1 093E-08
0 08	3 768E-08	3 768E 08	3 768E 08	3 774E 08	3 767E 08
0 084	1 299E-07	1 299E-07	1 299E-07	1 301E-07	1 299E-07
0 088	4 480E-07	4 480E-07	4 481E-07	4 488E-07	4 480E-07
0 092	1 545E-06	1 545E-06	1 545E-06	1 548E-06	1 545E-06
0 096	5 328E-06	5 328E-06	5 329E-06	5 336E-06	5 327E-06
0 1	1 836E-05	1 838E-05	1 840E-05	1 842E-05	N/A
0 12	1 731E-05	1 732E-05	1 733E-05	1 736E-05	N/A
0 14	1 630E-05	1 631E-05	1 633E 05	1 635E 05	N/A
0 16	1 535E-05	1 536E-05	1 537E-05	1 540E-05	
0 18	1 446E-05	1 447E-05	1 448E-05	1 450E 05	
0 2	1 362E-05	1 362E-05	1 364E 05	1 366E 05	

Table 7 4 The simulation of $dq_{C(L)}/dt$ using the implicit trapezoidal method by four different sweeping time resolutions TR/800, TR/400 TR/200 and TR/100 where TR (= τ_{tr}/10) is the rise time for the voltage ramp input The time scale is normalized to τ_{tr}

electrons) Therefore, the current through L is proportional to n/s or $\int n\,dt$ for very short times This means that minority carrier charge must accumulate before current can flow which is consistent with transport physics This aspect of the physics is missing from the NC model which contains only derivatives of minority carrier charge

7 4 3 Comparison of the Implicit Trapezoidal Method and the Gear-2 Method

Substituting $v_{BE}(t) = a\,t$ into (7 13) I obtain an analytical solution of $dq_{C(L)}/dt$ for a uniformly doped base layer

$$i(t) = \frac{dq_{C(L)}}{dt} = \frac{A\,q\,D_n\,n_i^2}{N_A(0)\,X_B}\,\frac{a}{kT/q}\,e^{\left[\frac{\frac{a\,t}{kT/q}}{\frac{1}{r_D}+\frac{a}{kT/q}}\right]}\,e^{-t/r_D} \qquad (7\ 25)$$

Eq (7 25) is used as the standard for comparison From Table 7 3 and Table 7 4 UDCS simulations of $dq_{C(L)}/dt$ is accurate

SLICE uses the implicit trapezoidal method as the default because it has a smaller local truncation error (LTE) The ratio of local truncation errors of the implicit trapezoidal method and the Gear 2 method is about 3 8 [Nag75] In some cases, however, Gear-2 may give accurate answers where the implicit trapezoidal method does not

I now study the local truncation error between the implicit trapezoidal method and the Gear-2 method for the $dq_{C(L)}/dt$ for $t < T_R$ The analytical solution is used as a standard for comparison From Table 7 3 and 7 4, I observe that the Gear-2 method requires about twice the sweeping-time resolution to achieve the accuracy of the

implicit trapezoidal method This observation is consistent with the 3 8 ratio mention above [Neg75] However the error of the simulation result of the Gear-2 method is acceptable

As a summary, the implicit trapezoidal method has a smaller local truncation error than that of the Gear-2 method The implicit trapezoidal method provokes oscillation if the controlling voltage of a current source abruptly levels off The abrupt levelling-off of the controlling voltage rarely happens in practical bipolar transistors however because of the presence of base resistance Thus, even though the excitation input voltage levels off suddenly, the base resistance causes the internal v_{gg} to slow down

7 5 Discussion

In this section, I compare the output collector currents, i_C predicted by the L and NC models (Fig 7 1) Again the analytical solution will be used as the standard for comparison

The analytical solution for a uniformly doped base for $dq_{C(NC)}/dt$ and $dq_{C(L)}/dt$ are given in (7 8) and (7 25) respectively Since $|i_C| = |i_{QS}| - |dq_C/dt|$, for the NC model the collector current, i_C, is

$$i_{C(NC)}(t) = \frac{A\, q\, D_n\, n_i^2}{N_A(0)\, X_B} \left[e^{\frac{s\, c}{kT/q}} - 1 - r_D\, e^{\frac{s\, c}{kT/q}} \frac{s}{kT/q} \right] \qquad (7\ 26)$$

The collector current predicted by the NC model is negative if

$$e/(kT/q) > 1/r_D \qquad (7\ 27)$$

For many modern-day transistors $\tau_{tr} \approx 1$ ns typically, and thus the slope of the input voltage must be less than $3\ 10^7$ V/s for $i_C > 0$

As a contrast, for the L model,

$$i_{C(L)}(t) = -\frac{A\, q\, D_n\, n_i^2}{N_A(0)\, X_g}\left[e^{\frac{\phi\, t}{kT/q}} - 1 - \frac{e^{\frac{\phi\, t}{kT/q}} - e^{-t/\tau_D}}{1 + \frac{kT/q}{\phi\, \tau_D}}\right] \qquad (7\ 27)$$

The collector current predicted by the L model is positive for all ϕ and τ_D

The possible negative i_C predicted by the NC model is physically incorrect. For the intrinsic region of a bipolar transistor, i_C cannot be negative for any kind of turn-on voltage. This negative i_C is related to the fundamental limitation of the NC model, which was discussed in Chapter Six, Section 5.

7 6 Application: A Simple Transistor Inverter

I shall perform SLICE/UDCS simulations of the turn-on transients of the single-transistor inverter circuit (Fig. 7 2). The large signal Gummel-Poon, SC, and the L models are used in the simulations. This computer experiment resembles that in Fig. 2 of Foesum and Vesterraghavan [Fes86], in which they illustrated that the NQS delay, accounted for in the non-quasi-static models but not in the Gummel-Poon model, is reflected by the horizontal displacement among the simulated transient responses. In their simulation, they use one value of the input resistance, $R_{in} = 1000$ ohms

In this section I repeat their simulation in the sense that I use the same model parameters. I extend their simulation by using three different values of R_{in}: 50 ohms, 1000 ohms, and 20 000 ohms. Other differences are:

(1) I include the large-signal L model and the large-signal NC model in the simulation in addition to the Gummel-Poon model,

(2) the input voltage is a ramp from 0 V to 0.75 V with a rise time of half of the base transit time (see the inset of Fig. 7.2), instead of a voltage step with zero rise time from 0.6 V to 0.75 V [Fos86],

(3) the time scale used in the simulation is normalized to the base transit time, τ_{tr}.

In this computer simulation, I verify a fundamental limitation of the NC model, which was discussed in Chapter Six.

Small R_{in}. Here I use R_{in} = 50 ohms. The collector currents predicted by the Gummel-Poon (GP), the large-signal NC and the large-signal L models are shown in Fig. 7.2, in which I observe that:

(1) the non-quasi-static delay (the horizontal distance between the i_C's predicted by the GP and L model) is significant and is about one third of the base transit time, τ_{tr}.

(2) the NC model predicts a very large negative overshoot in the collector current. This is consistent with the analytical solutions discussed in Section 7.4. This negative overshoot is related to a fundamental limitation of the NC model (see Chapter Six).

Large R_{in}. Here I use R_{in} = 20,000 ohms. The simulation results are shown in Fig. 7.3, in which I observe that: (1) although the non-quasi-static delay is still about one third of the transit time,

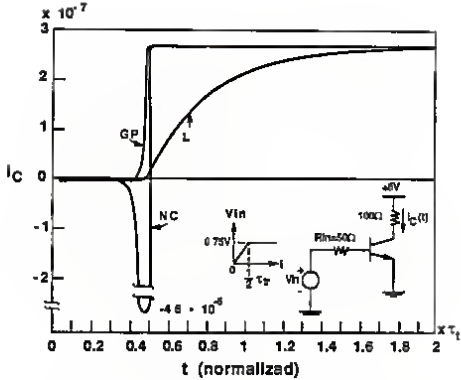


Fig. 7.2 SLICE/SPICE simulations of the turn-on transient of the single-transistor inverter circuit shown in the inset (R_{in} = 50 ohms). The large signal L and NC models and the Gummel-Poon model are used in the simulations. The NQS delay is reflected by the horizontal displacement among the I_C's predicted by the GP and the L models. The time scale is normalized to τ_{tr}

it is however not as significant as that for small R_{in} (Fig 7 2) The difference in i_C between the GP model and the non quasi-static model is about 10% (2) the NC model predicts very little negative overshoot in the collector currents In fact very little difference between i_C's is predicted by the NC and the L models

Intermediate R_{in} Here I use R_{in} = 1000 ohms The simulation results are shown in Fig 7 4, in which I observe that
(1) the non-quasi-static delay is about one third of the transit time
(2) The NC model predicts negative overshoot in the collector currents It is interesting to note that according to the area approach the area difference of the collector currents predicted by the GP and NC models and by the GP and L models are equivalent

Conclusion. Non quasi-static effects are less important for base current drive (which is equivalent to a voltage drive with a very large R_{in}) If I have a voltage drive with small R_{in}, the non quasi-static delay and the negative overshoot in i_C predicted by the NC model are significant

7 7 Summary

I have implemented the large-signal L and NC models using the UDCS approach Detailed documentation provided in this chapter will be useful for implementing the L and NC models directly into SPICE2 by rewriting the code for more efficient simulation

I compared two different numerical methods that transform the differential equation to difference equations (in time) the implicit-trapezoidal method and the Gear-2 method For the NC model, the implicit-trapezoidal method led to severe oscillations in the

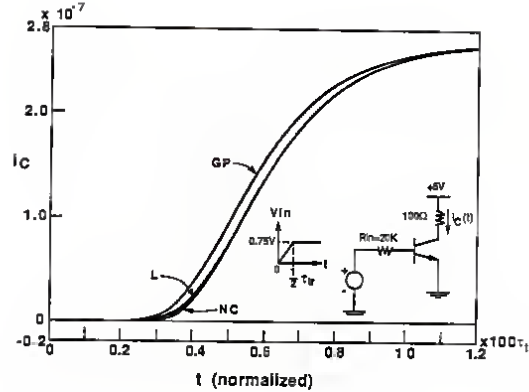


Fig 7 3 SLICE/SPICE simulations of the turn-on transient of the single transistor inverter circuit shown in the inset (R_{in} = 20 000 ohms) The large signal L and NC models and the Gummel-Poon model are used in the simulations The NQS delay is reflected by the horizontal displacement among the curves The time scale is normalized to τ_{tr}

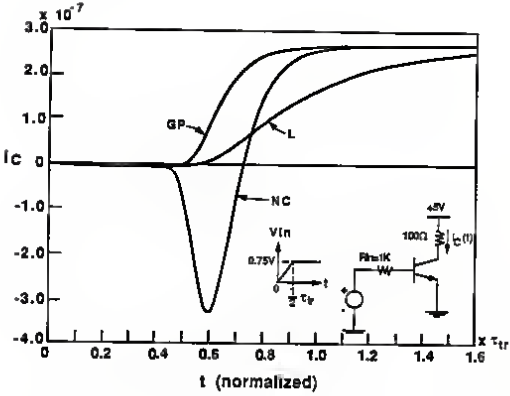


Fig. 7.4 SLICE/SPICE simulations of the turn-on transient of the single-transistor inverter circuit shown in the inset (R_{in} = 1000 ohms). The large-signal L and NC models and the Gummel-Poon model are used in the simulations. The NQS delay is reflected by the horizontal displacement among the i_C's predicted by the GP and the L models. The time scale is normalized to τ_{tr}

capacitive current sources (including the transcapacitance current source) in the model. The reasons for this oscillation were determined. To rectify this problem, I used the Gear-2 method instead of the implicit trapezoidal method. For the L model, both the implicit-trapezoidal and the Gear-2 methods worked well. For the simplest case of homogeneously doped base layers, the accuracy of the UDCS simulation results were assessed by comparison with analytical solutions, and were shown to be excellent. The simulation of the L model and the NC model will require comparable CPU time because no integration subroutine is required for simulating $dq_{C(L)}/dt$. I then perform a computer simulation of a single-transistor inverter with different values of R_{in}. For a smaller R_{in}, the NC model shows an undesirable negative overshoot in the output collector current of a step voltage input.

CHAPTER EIGHT
A SELF-CONSISTENCY TEST FOR DEVICE MODELS IN TRANSIENT COMPUTER SIMULATION OF ELECTRONIC CIRCUITS

Most device models used in electronic circuit simulators such as SPICE are based on the quasi-static approximation. At high frequencies or for transients this approximation introduces error. But presently no method exists for systematically assessing these devices in a circuit for which quasi-static models must be replaced by more accurate (non-quasi-static) models if accurate circuit simulations are to be obtained. The purpose of this chapter is to develop and demonstrate such a method.

In this development, I first review the quasi-static approximation and its role in idealized and extended charge control models. A test for the self-consistent validity of each quasi-static model in a circuit simulation is then developed. Only those quasi-static models in a particular circuit that fail the test need to be replaced by non-quasi-static models.

The test applies to MOS as well as bipolar circuits. To illustrate the test, I apply it to several simple bipolar circuits.

8.1 Introduction

As discussed in Chapter Two, from a fundamental standpoint, the analysis of such semiconductor devices as bipolar and MOS transistors

is based on a set of differential equations [Shs49], which expresses relationships among currents, mobile-carrier concentrations, and electric potential. Analytic solution of these equations, without approximations, in the full generality needed to describe dynamic response, has proven to be intractable. Hence, to simplify treatment of these basic equations, various methods of approximation have evolved.

Among these is the quasi-static approximation. As was noted in earlier chapters, in this approximation one first omits all time-dependent terms appearing in the basic equations. Analytic solution then becomes possible if various additional approximations are employed. This solution yields expressions relating the currents and charge components within the device to the boundary conditions, for example, to the terminal voltages. These expressions apply only if the boundary conditions are time invariant. To obtain an approximate solution for time-varying boundary conditions, one makes the quasi-static approximation that the expressions hold in general, despite departures from the dc steady state. Most device models used in the computer-aided analysis of bipolar and MOS circuits depend for their validity on the quasi-static approximation.

The self-consistency test to be developed applies both to bipolar and MOS transistors, and to other devices as well. Use of the concept of the ideal charge-control model, extended appropriately, enables a development in this generality. Using the extended ideal charge control model, I then develop the test for self-consistency and unify alternative views of its meaning. I emphasize the extrinsic as well as

the intrinsic device and various nonidealities that distinguish the behavior of actual devices from that of the ideal charge-control model

As concrete examples I present the results of applying the quasi static self-consistency test to bipolar devices In a practical inverter circuit Marked violations of the quasi-static approximation are shown to occur An estimate is made of the effect these device model violations have on predicting system parameters such as the turn ON and turn OFF propagation delays This estimate involves the use of non-quasi-static bipolar models

§ 2 Self Consistency

In conventional device analysis the device is divided into extrinsic and intrinsic parts as Fig § 1 illustrates The self consistency test described here applies to both of these parts I begin with the intrinsic device

§ 2 1 Self Consistency from Extended Ideal Charge-Control Model for the Intrinsic Device

Charge-control analysis was originally applied only to photoconductive devices (Ros54) and to junction transistors [Ros57] In 1959 however, Johnson and Rose [Joh59] and Middlebrook [Mid59] introduced charge control as a means to unify characterization of many different kinds of electronic devices I exploit this unification here, together with the property that charge control is rooted in the quasi static approximation

Earlier work will guide our treatment Previous presentations and applications of charge control do not satisfy our needs for assessing

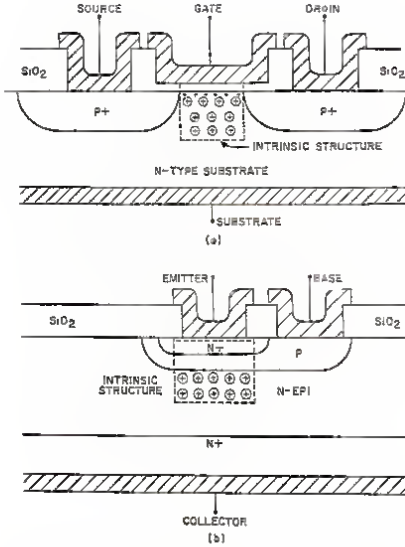


Fig. 8-1 Illustrating the division into intrinsic and extrinsic devices. The dotted boxes enclose the intrinsic structures (Not to scale)
(a) Power MOSFET, (b) power bipolar transistor

self consistency, however, partly because most earlier workers concentrated on small-signal, amplifying applications, partly because they restricted consideration only to the intrinsic device, and partly for other reasons that will become plain as our treatment progresses.

Following others [Joh59, Mid59] I begin by considering the general three-terminal electronic device pictured in Figure E 2. This is a charge-control device, which means that the current i_{21} flowing between terminals 1 and 2 is controlled by the charge q_3 placed at terminal 3. I model the device using an ideal charge-control model defined by the following properties:

(a) The current i_{21} flowing between terminal 1 (emitter or source, for example) and terminal 2 (collector or drain, for example) is the ratio of the controlled mobile carrier charge q_{21} to the mean static transit time t_{21} needed for these carriers to move between these two terminals. Hence,

$$i_{21} = q_{21}/t_{21} \qquad \text{(E 1)}$$

(b) The controlling charge q_3 associated with terminal 3 (base or gate, for example) exercises total control over the controlled charge according to

$$dq_{21} = -dq_3 \qquad \text{(E 2)}$$

(c) The current i_3 flows only to change the charge q_3 placed there.

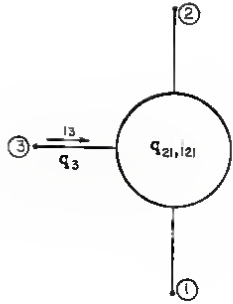


Fig. 8.2 General charge control device

Thus

$$i_3 = \dot{q}_3 \qquad (8.3)$$

where the dot notation designates time differentiation.

In all devices, property (a) is an idealization valid in a strict sense only in the dc steady state, but extrapolated to apply at each instant during transients by use of the *quasi-static approximation*. Although in previous works [Bee57, Ree54, Mil59] the use of the quasi-static approximation was not explicitly discussed, this approximation was implicitly assumed. (In fact, the term quasi-static approximation was first seen in Hamilton and coworkers [Ham63].) Property (b) is an idealization for most devices in that q_3 only partially controls q_{21}. For example, in an MOS transistor, some of the flux lines starting from gate charge may end on depletion charge rather than on the charge of mobile carriers in the channel. As another example, in a short-channel MOS transistor, the two-dimensional (or three-dimensional) configuration of flux lines in the insulator and in the semiconductor may invalidate the full control expressed in (8.2). Similar violations of (8.2) occur in bipolar transistors and other devices. Property (c) is an excellent approximation for some devices, an example is the MOS transistor. For the bipolar transistor, however, it is an idealization because the base fails to retain all of the charge q_3 placed at its terminal. Net recombination occurring within the device or at its surfaces or contacts, for instance, requires a component of i_3 in addition to that expressed in property (c).

For the purposes of this work however, these nonidealities are inconsequential As I shall see for the bipolar and the MOS transistor formulation of a test for self-consistency of the quasi-static approximation can be based on straightforward extensions of the ideal charge control model defined above

To specify the network representation of the ideal charge control model, consider first the simple case in which terminal 2 acts as the sole collector of the transport current i_{21} This corresponds approximately to forward-active operation for the bipolar transistor and to pinch-off operation for the MOSFET For this case, the properties defined by Equations (8 1) - (8 3) imply the network representation of Fig 8 3(a) in which chain rule differentiation applied to Eq (8 3) gives for the capacitance

$$C_{31} = \frac{\delta q_3}{\delta v_{31}}\bigg|_{v_{32}} = -\frac{\delta q_{21}}{\delta v_{31}}\bigg|_{v_{32}} \tag{8 4}$$

Fig 8 3(a) represents the large- and small-signal transient response, in contrast to the small-signal representations of Johnson and Rose [Joh59] and Middlebrook [Mid59]

Some circuit applications reverse the roles of terminals 1 and 2 with terminal 1 becoming the sole collector of the current i_{21} as in the inverse-active operation of the bipolar transistor This corresponds to the network representation of Fig 8 3(b), with

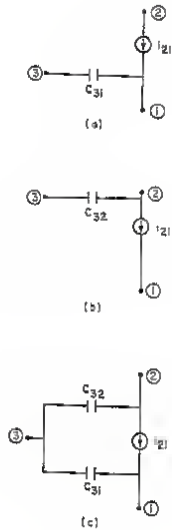


Fig 8 3 Illustrating the development of the general network representation, shown in (c) of the extended ideal charge control model It approximates the behavior of the *intrinsic* part of bipolar transistors MOSFETs and other charge control devices

$$C_{32} = \left.\frac{\partial q_3}{\partial v_{32}}\right|_{v_{31}} = \left.\frac{\partial q_{21}}{\partial v_{32}}\right|_{v_{31}} \qquad (\S\ 5)$$

In the most general case, terminals 1 and 2 both collect and emit the transport current i_{21}, as in the saturated mode of the bipolar transistor and the non-pinch-off mode of the MOSFET. Thus the general network representation of the ideal charge-control model *extended* to apply for large-signal applications and to account for the dual roles of terminals 1 and 2, derives from a parallel combination of Figs. § 3(a) and § 3(b). The general network representation appears in Fig. § 3(c).[1]

In this network, the circuit elements are all functions of the device make up and the terminal voltages and currents. The particular functional dependencies that apply result from the details of the physical modeling used to characterize the charge q_{21} and the transit time t_{21}. For example, the Ebers-Moll model [Ebe54] implies

$$q_{21} = \frac{q\, n_i^2\, X_B}{2N_{AA}} \exp(qv_{EB}/kT) \text{ and } t_{21} = X_B^2/2D_B \qquad (\S\ 6)$$

for the bipolar transistor in the forward-active mode, and the simple square-law characterization [Sah64, Lin72] for light substrate doping

---

[1] Rather than the plausibility argument given, one can derive the network representation of Fig. § 3(c) by more strict reasoning. This representation follows from chain-rule differentiation of Eq. (§ 3) joined with the implication of Eq. (§ 2) that all flux lines of Coulomb force associated with the control of q_{21} by q_3 must contact terminal 3 (name link terminals 1 and 2).

implies

$$q_{21} \approx (2/3)C_0 Z L v_{GS} \text{ and } t_{21} \approx (4/3)L^2/\mu v_{GS} \qquad (8.7)$$

for the MOSFET in pinch-off operation. Here C_0 denotes the gate oxide capacitance per unit area, μ, the mobility in the channel, v_{GS} the gate source voltage, L the channel length, and Z its width.

Self consistency criteria. As was noted, the ideal charge-control model of Fig. 8.3(c) is based on the quasi-static approximation. Because this approximation is exact in the dc steady state, I take the following as the criterion for self-consistency: the configuration of current flow in the network representation during transients must correspond to only a small perturbation of the configuration existing in the dc steady state. Hence, from Fig. 8.3(c), the criterion is

$$\frac{|C_{31} v_{31}| + |C_{32} v_{32}|}{|I_{21}|} \ll 1, \qquad (8.8)$$

or, in more general notation, the merit of which will be seen later,

$$\frac{\sum_{jk} |C_{jk} v_{jk}|}{|q_{21}/t_{21}|} \ll 1. \qquad (8.9)$$

Expression (8.8) and (8.9) constitute necessary and sufficient conditions for self consistency. As we shall see, they form the basis

for a self-consistency test that can be implemented in computer programs for circuit analysis

View from another standpoint I see that a consistent but less demanding necessary condition is

$$\frac{|\dot{q}_3|}{|q_{21}/\tau_{21}|} = \frac{|\dot{q}_{21}|}{|q_{21}/\tau_{21}|} \ll 1, \tag{8-10}$$

Let the applied excitation make q_{21} vary in proportion to $e^{i\omega t}$ Then (8-10) becomes

$$\omega \tau_{21} \ll 1 \tag{8-11}$$

which implies self consistency with the quasi-static approximation provided the period of the applied signal greatly exceeds the transit time of the mobile carriers If this holds, the mobile carriers can spatially distribute themselves fast enough to assume essentially the same relation to the boundary conditions as exists in the dc steady state

These self-consistency criteria follow from the assertion that the quasi static approximation is exact as $\omega \to 0$ Despite the intuitive appeal of this assertion and although it has been advanced by others, the observations of Cherry [Che71] about capacitance indicate that it is not strictly true From a practical standpoint, no change in our conclusions results, however, from regarding the quasi static approximation as exact in the dc steady state The device models I

consider consist of controlled sources and resistors which represent the flow of conduction current, and other circuit elements usually capacitors which represent capacitive effects That the magnitude of the capacitors calculated by quasi static analysis may be in error even for arbitrarily low frequency of excitation, does not affect our conclusions, because at such frequencies the currents flowing through these capacitors are most often negligible relative to the conduction currents That the magnitude of these quasi-static (or charge-control) capacitors may be in error at higher frequencies contained in the excitation likewise does not affect our conclusions Rather it influences only the strength of the inequalities appearing in such expressions as (8 9)

Effect of nonidealities on self-consistency criteria The criteria expressed in (8 8) and (8 9) are based on the ideal charge-control model which provides only a first approximation to the performance of actual devices Nonetheless these criteria stay valid in the presence of commonly occurring nonidealities as I shall now show

There are several kinds of such nonidealities Consider first the kind that arises from net recombination or generation occurring within the intrinsic device The drain current of the MOS transistor, for example, can increase with increasing drain voltage because of avalanche multiplication of mobile carriers--a form of net generation One must account for this nonideality in the modeling of the mobile charge, q_{21} and of the (effective) transit time, t_{21}, but the network representation remains that of the ideal charge control model shown in

Fig. 8.3(c). Hence the self consistency criteria of (8.8) and (8.9) remains valid for a three terminal model of the MOSFET despite the occurrence of avalanche multiplication. The same conclusions hold for net recombination or generation in a MOSFET that arises from other processes, such as thermal emission or capture of carriers, provided the resultant current flowing in the dc steady state to the substrate is negligible.

In contrast, net generation or recombination in a bipolar transistor requires the addition of controlled current sources in parallel with each capacitor in the network representation of Fig. 8.3(c). The time constants normally associated with these sources will differ markedly from the transit time for mobile carriers. From a practical viewpoint, however, except for operation near avalanche breakdown, the additional sources will carry small current relative to that of source i_{21}, and the self consistency criteria for the quasi-static approximation remain as I have stated them.

In a previous section, I discussed a second kind of nonideality that arises because the controlling charge q_3 fails to exercise full control over the controlled charge q_{21}. One must attend to this nonideality in modeling the charges and the transit time; but, again, it affects only the functional dependencies of the circuit elements in the network representation of Fig. 8.3(c), not the representation itself. Hence, this kind of nonideality also requires no change in the self-consistency criteria for a three terminal model of device behavior.

In MOS transistors there is another kind of nonideality of central importance Many large signal circuit applications cause the potential of the substrate to change with time rapidly sometimes relative to the potentials of the source or gate or drain Hence the potential or charge, associated with the substrate can influence the charge q_{21} of mobile carriers in the channel This necessitates an extension of our previous view of a charge control device and of the ideal charge control model which I will exclude from this work Tsividis [Tsi87] gives a careful, detailed treatment of the large-signal dynamic modeling of the four-terminal MOSFET

§ 2 2 Model for the Extrinsic Device

The division shown in Fig 8 1 determines that the intrinsic part is a charge-control device whose behavior is approximated by properties of Eqs (3 1) through (3 3) given in the previous section On the other hand this same division determines that the behavior of the extrinsic part will not depend on charge control For example the number of mobile carriers contained within the extrinsic part of the transistor (such as the source or drain islands of a power MOS transistor) is fixed to a good approximation by the resistivity of these regions independently of the charges associated with the terminals bordering the regions Hence the modeling of the extrinsic device follows a different scheme than that described for the intrinsic device which requires approximate solutions for the dc steady state of the basic differential equations The scheme for the extrinsic structure depends on intuitive reasoning and much of the modeling is

dens by inspection [Lin71b] As we shall see however the self-consistency criterion for the extrinsic part greatly resembles the criteria I have already developed

Because treatment of the extrinsic part cannot be unified in the framework of the charge control model extrinsic network representation depends strongly on the details of device structure Even within one category of device the MOSFET, for example, different structures may require different network representations, and typically the representations for MOS and bipolar transistors will be highly dissimilar

Consider first the dc steady state The dominant process occurring within the extrinsic transistor is the current conduction through resistive material to the terminals I model this by the three resistors connecting intrinsic terminals to actual terminals shown in Fig 8 4

Consider now the transient response In addition to conduction capacitive effects also occur For a MOS transistor, the capacitive effects arise from the overlap of the gate metal above the source and drain islands and from the pn junctions of these islands Because the capacitive effects are distributed RC transmission lines, in which nonlinear capacitances represent the pn junctions, offer a refined approximation of the behavior From a computational standpoint however, they pose such severe difficulties in circuit analysis that one reverts to coarser approximations [Lin71b], for instance, the lumped circuit of Fig 8 4 consisting of one resistor and two capacitors for either MOSFET or BJT

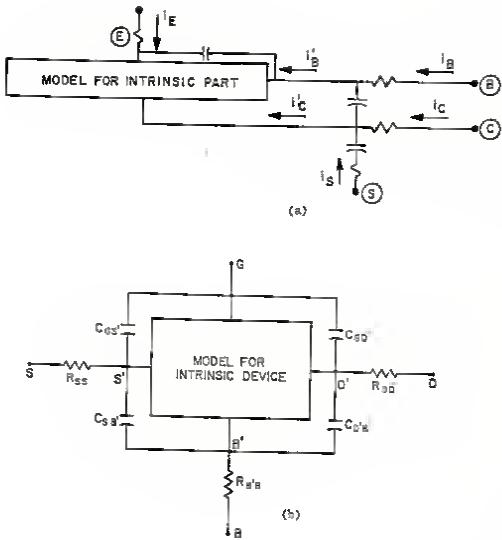


Fig 8.4 Possible network representation of the extrinsic part of the transistor pictured in Fig. 8.1. (a) BJT, (b) MOSFET

Thus the model of Fig. 8.4 for the extrinsic device comes from approximating a distributed network by a lumped network. This is a quasi-static approximation, which is exact in the dc steady state. Therefore, as for the intrinsic part, I take the following as the self-consistency criterion for the network representation of the extrinsic device: the configuration of current flow in the network representation during transients must correspond to only a small perturbation of the configuration existing in the dc steady state. Hence, from Fig. 8.4(b), the criterion is

$$\frac{|C_{GS'}\dot{v}_{GS'}| + |C_{S'g'}\dot{v}_{S'g'}|}{|v_{SS'}/R_{SS'}|} \ll 1 \qquad (8.12)$$

for the network representation of the source island. An analogous criterion applies for the drain island. If more circuit elements are used to approximate the distributed line, the same criterion holds with (8.12) now interpreted as applying to each three-element section of the network representation.

8.2.3 Self-Consistency Test for Large-Signal Excitations

A circuit whose behavior is to be simulated by computer may involve many devices, each of which may have a network representation composed of the models described here for the extrinsic and intrinsic parts. Typical programs simulate the currents in each branch of the network representations at each time step. Inserting the values of these currents in (8.9) for the intrinsic device and in (8.12), or a

suitable counterpart for the extrinsic device then constitutes a test for the self consistency of the quasi static approximation and of the models that are based on it. Note that the intrinsic and extrinsic parts of a device model must each be tested for obedience to the self consistency criteria.

Implementation of the self-consistency test in BJT (bipolar junction transistor) circuit simulation will be treated in Section 8.4 together with discussions of the violations of self-consistency from a physical standpoint and of their influence on the system parameters of BJT circuits.

8.2.4 Self Consistency Test for Small-Signal Sinusoidal Excitations

Equation (8.11) provides a simple, intuitively appealing test for small-signal sinusoidal excitation that avoids the details discussed in Section 8.2.3.

8.3 Model Updating

If a particular device model fails to pass the test in a circuit simulation a more complex model is needed to assure self-consistency with the quasi static approximation.[2] To help determine the degree of added complexity that is necessary one can again employ the self-consistency test extended by the following strategy. Recall that the

[2] It is possible for device models to radically fail the self consistency test and yet to serve adequately in the simulation of such system parameters as the turn off propagation delay time. This issue is explored later.

test is to be applied individually to the extrinsic and intrinsic models of the device.

Extrinsic model. Consider first the model for the extrinsic structure. If the simple three-element model of Figure 4 disobeys the inequality expressed in (8.12), and thus violates self consistency, an additional three-element section can be added to give a better approximation of the distributed nature of the extrinsic structure. Self consistency is then tested again, with (8.12), now interpreted as applying for each section of the extrinsic model. This procedure is repeated until self-consistency ultimately prevails in the extrinsic model. That it will prevail is assured because the capacitance and resistance per section both diminish as more sections are added. Ultimately, as more sections are added, the resistive current in each section will much exceed the capacitive current and (8.12) will be obeyed. Notice that the self-consistency test will determine how many sections must be added to cause this to occur.

Adding complexity to the extrinsic model by itself tends to assure self-consistency of the model for the intrinsic structure. This occurs because the more complex the extrinsic model, the more effectively it filters out any higher frequencies contained in the excitation at the device terminals, preventing them from reaching the terminals of the intrinsic model. Hence, in accord with the alternative criterion for self-consistency expressed in (8.11), a detailed model for the extrinsic structure tends to enhance the self-consistency of the model for the intrinsic structure.

Despite the complexity of the extrinsic model the model for the intrinsic device may still violate self-consistency by failing to obey inequalities (8.9). To remedy this, just as for the extrinsic model one can complicate the network representation to give a better lumped network approximation of the distributed nature of the intrinsic structure.

Multisection model. One strategy for model updating would require generating multi section models for the intrinsic device, similar to those for the extrinsic structure discussed immediately above and in Section 8.2.2. Various techniques [Ham71, Sah70, Sah71] might be used to obtain such models. The basic idea here is simple. Decreasing the size of the region being modeled ultimately brings the boundaries close enough together to enable mobile carriers within the region to distribute themselves so rapidly that they assume essentially the same relationship to the boundary conditions during transients as exists in the dc steady state. Thus, use of enough sections in an intrinsic model ultimately yields self-consistency with the quasi-static approximation. To determine the degree of complexity needed, one could employ the test implied in (8.5), extended in interpretation to apply section by section.

Non quasi-static models. Several (non quasi static) models for the BJT and MOSFET that do not require the use of the quasi-static approximation have been proposed. For non quasi-static MOSFET models see [Arr77, Wat78, Oh80, Tei87]. For non quasi static BJT models, see Chapter 51. The advantage of using non-quasi-static models rather than the multisection quasi-static model is that generally non quasi-

static models require fewer circuit elements to achieve the same accuracy. Note, however, that the quasi static test is inapplicable to non-quasi-static models.

8.4 BJT Inverter Case Study

The results of applying the quasi static self-consistency test to a practical digital BJT circuit are presented below. The BJT device model used in the simulation is a quasi-static (QS) model (Fig. 5.3a) which is the Gummel-Poon model used in the SPICE2 circuit simulation program. To provide a quantitative estimate of the error in switching times that this quasi-static BJT model provokes, I can compare against simulations based on two non-quasi-static (NQS) models: the large-signal L and NC models. This large-signal BJT models are implemented using the UDCS technique described in Chapter Seven.

Although I use both the L and NC models in the case study, I will only show the simulation results of the non-quasi static L model to avoid repetition of the work in Chapter Seven especially Section 7.6, in which I studied a transistor inverter with different value of the input resistance using the Gummel-Poon, the L and the NC models. In Section 7.6, I demonstrated that the NC model predicted a large negative overshoot in the collector current for a voltage drive with small input resistances (50 Ω and 1 KΩ). For a current drive, which is equivalent to a voltage drive with a very large input resistance, the NC models showed little negative overshoot. Non-quasi-static effects are less important for a current drive. For details, refer to Section 7.6 especially Fig. 7.2, 7.3 and 7.4.

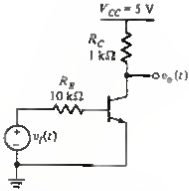
Application of the quasi-static approximation to the intrinsic device requires that the sum of the absolute value of the currents through the intrinsic capacitors be much less than the average value of the transport current. Therefore a test is proposed such that a violation occurs when the quasi-static ratio, QSR, is greater than unity, i.e.

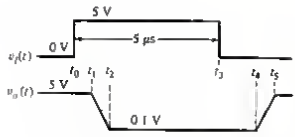
$$QSR = \frac{|i_{CBE}| + |i_{CBC}|}{i_{QS}} > 1 \qquad (8.13)$$

where i_{CBE} and i_{CBC} are the currents flowing through the intrinsic capacitors associated with the emitter and collector, respectively.

Case study-1. The circuit used in this case study is a basic element of most digital systems, the logic inverter. A simple but practical configuration of a bipolar transistor logic inverter is shown in Fig. 6.1 in [Hod83] and repeated here as Fig. 8.5(a), along with transistor data, which are typical of a low power n/p/n switching transistor [Hod83]. In Fig. 8.5(b), I list the switching times of interest. A detailed explanation of the physical mechanisms occurring within the transistor during each phase of the transient is given in Hodge and Jackson [Hod83].

The output voltage of the inverter circuit (Fig. 8.5a) predicted by the quasi-static model is shown in Fig. 8.6. The input voltage is a rectangular pulse that at 1 ns abruptly changes from 0 to 5 V to turn the transistor ON and changes from 5 V to 0 at t = 21 ns to turn the





$V_{BE(on)} = 0.7$ V	$\tau_F = 0.2$ ns
$V_{BE(sat)} = 0.8$ V	$\tau_{BF} = 14$ ns
$V_{CE(sat)} = 0.1$ V	$\tau_S = 20$ ns
$C_{je0} = 0.30$ pF	$C_{je0} = 0.15$ pF
$\phi_e = 0.9$ V	$\phi_c = 0.7$ V
$m_e = 0.50$	$m_c = 0.33$

(a)

Delay time	$= t_d = t_1 - t_0$
Fall time	$= t_f = t_2 - t_1$
Saturation time	$= t_s = t_4 - t_3$
Rise time	$= t_r = t_5 - t_4$

(b)

Fig. 8.5 A simple but practical configuration of a bipolar transistor logic inverter. After Hodge and Jackson [Hod83]. (a) BJT inverter; (b) input and output voltage waveforms.

transistor OFF. The switching times predicted by the QS model and the NQS model are tabulated in Table 8.1.

One method of evaluating the violations is to examine (1) the maximum value of QSR, (2) the time interval when QSR > 1, and (3) a weighting QSR factor during the switching time when QSR > 1, as shown in Eq. 8.14. Equation (8.14) is illustrated for t_d, but can be applied to t_f, t_a and t_r (defined in Fig. 8.5b) also.

$$\text{weighting QSR } (t_d) = \frac{\int_{t_0}^{t_1} \text{QSR}\, dt}{t_1 - t_0} \qquad (8.14)$$

In Fig. 8.6, the QSR indicates that the quasi-static approximation (QSA) is valid for most of the transient except for initial part of the turn-on transient. The validity of the use of the QSA is confirmed from the negligible error in the switching times, t_f, t_a, and t_r predicted by the QS model (Table 8.1). Since the quasi-static approximation is violated during the initial part of the turn-on transient (1 ns < t < 1 ns + t_d), we expect that the delay time, t_d, would be affected. There is, however, only a 4% error in t_d predicted by the QS model. This is because for a typical low-power switching transistor the turn-on transient is dominated by extrinsic elements such as the base-collector junction capacitance (C_{jc}). This dominance of C_{jc} is obvious in the small peak in the output voltage, observable in Fig. 8.6 slightly after t = 1 ns, which is caused by the feedthrough of the input pulse via C_{jc}, while the transistor is still off. I have

R_g	switching times	QS model ns	NQS L model ns	Error	QSR max	QSR ns	QSR weighting
10K	t_d	1 10	1 15	4%	7 1	0 83	2 1
	t_f	3 28	3 29	< 1%	< 1		
	t_s	17 28	17 38	< 1%	< 1		
	t_r	11 32	11 51	2%	< 1		
2 2K	t_d	0 29	0 30	3%	6 0	0 29	2 5
	t_f	0 70	0 73	4%	3 0	0 50	1 1
	t_s	26 06	26 09	< 1%	< 1		
	t_r	3 21	3 25	1%	< 1		
22K	t_d	2 34	2 38	2%	3 1	0 45	0 4
	t_f	8 27	8 29	< 1%	< 1		
	t_s	9 52	9 56	2%	< 1		
	t_r	18 17	18 22	< 1%	< 1		

TABLE 8 1 Switching times of the inverter (Fig 8 5a) with R_g = 10K 2 2K and 22K The maximum value of QSR, the time interval when QSR > 1 and the weighting QSR factor Eq (8 14) are given

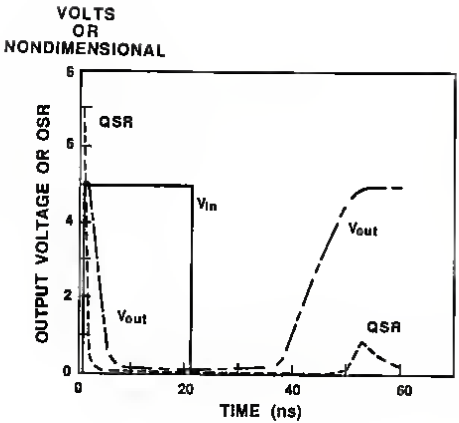


Fig 8 6 UDCS (SPICE) simulation of the transient response for the simple transistor inverter with R_B = 10K The quasi-static ratio (QSR) indicates that the quasi-static approximation is violated at 1 ns

used three different values of R_B (2.2K, 10K and 22K), the above observation applies to R_B = 2.2K and 22K also.

Case study-2: I replace the low-power switching transistor used in Fig. 8.5(a) by a power transistor and repeat the computer experiment. For a typical power transistor [Mil80, Gte84, Ant86], the non quasi-static effects will be more significant than those of a low power switching transistor because power transistors require thick quasi-neutral base layers. The major differences between the transistor in Fig. 8.5(a) and the power transistor are reflected by τ_F = 0.1 μs and τ_R = 10 μs for the power transistor.

Figure 8.7 and Table 8.2 show the transient responses and switching times, respectively, for the inverter circuit of Fig. 8.5(a) using a power transistor with different values of R_B. From Table 8.2 and Fig. 8.7, we observe the following:

(1) The L non-quasi-static model predicts as much as a 33% longer delay t_d than that predicted by the Gummel-Poon model.

(2) The absence of the small peak in the output voltage indicates that the effect of C_{jC} is not important.

(3) The smaller R_B is, the more the error in the switching times predicted by the QS model.

(4) The larger values of QSR and weighting QSR factor (8.14) observed correspond to larger errors in switching times predicted by the QS model.

(5) The quasi-static approximation is not violated in the saturation mode of operation.

(6) The quasi-static approximation is slightly violated in the recovery

R_B switch times		QS model x10^{-7} s	NQS L model x10^{-7} s	Error	QSR max	QSR x10^{-7}s	QSR weighting
10K	t_d	1 1	1 4	21%	9 5	1 1	3 0
	t_f	10 8	11 0	2%	< 1		
	t_s	54 9	55 2	< 1%	< 1		
	t_r	34 3	34 7	1%	< 1		
2 2K	t_d	0 2	0 3	33%	11	0 2	6 6
	t_f	2 6	2 9	10%	4	0 7	0 6
	t_s	82 8	83 5	< 1%	< 1		
	t_r	10 7	11 2	4%	< 1		
22K	t_d	2 8	3 0	13%	12	1 6	2 0
	t_f	28 8	28 7	< 1%	< 1		
	t_s	29 7	30 0	1%	< 1		
	t_r	53 2	53 3	< 1%	< 1		

TABLE 8-2 Switching times of the inverter (Fig 8 5a) with R_B = 10K 2 2K and 22K The BJT device used here is a typical high power devices [Mil80 Gre84 Ant58] The maximum value of QSR, the time interval when QSR > 1 and the weighting QSR factor, (§ 14), are given

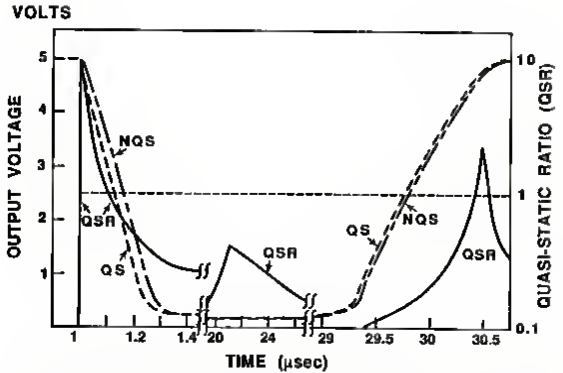


Fig. 8 7 UDCS (SPICE) simulation of the transient response for the simple transistor inverter with R_B = 2 2K The non-quasi-static effects are significant in this transistor a typical power transistor The transient responses for R_B = 10K and R_B = 22K are similar to that of R_B = 2 2K and are not shown The NQS simulation result is predicted by the L model

region [Hod83] which is the transient phase at $t > t_5$ shown in Fig 5 5 However since the final recovery time is not a parameter of interest to the digital IC designer [Hod83] the error in the final recovery time predicted by the QS model is not given

5 5 Discussion

The QSR test can be implemented in many circuit simulation programs It has been shown here that the quasi-static models used for power transistors exhibit violations of the quasi-static approximation and errors in predicting the delay time t_d

Application of the QSR test to determine the limit of model complexity is useful in choosing models for large circuit simulations For example application to a fifty-transistor circuit may reveal that only two or three device models are in violation Only these models need to be replaced by more accurate models

The self-consistency criteria derived here apply generality to all electronic devices whose intrinsic structure operates in accord with charge control The criteria hold for all regions of device operation and include consideration of the extrinsic part of the device In previous related discussions [Lin71b, Gro64, Gre69] the scope was more limited Those treated the intrinsic device only and concentrated only on the region of operation in which terminal 2 of Fig 5 2 acts as the sole collector of the controlled current that is, forward-active operation of the bipolar transistor and pinch-off operation of the MOSFET

In the present work nonidealities are defined as departures from the behavior predicted by the extended ideal charge control model set forth in Section E 2 2 Nonidealities receive careful attention Despite their existence, the following statement derived for the ideal charge control model holds generally the configuration of current flow in the network representation of large-signal device behavior during transients must correspond to only a small perturbation of the configuration existing in the dc steady state Otherwise the quasi static approximation, on which the network representation is based will be violated The statement applies individually to the representations of the intrinsic and the extrinsic part of a device

The general statement above is phrased mathematically in expressions (3 9) for the intrinsic structure and in (3 12) for the extrinsic structure It enables self consistency to be tested by comparing currents in various branches of the network representation Since these currents are routinely computed in any simulation of circuit behavior the test is easily implemented in circuit analysis programs

E 6 Summary

I proposed a test for the self consistent validity of the quasi static approximation The test applies to each device model in a circuit and is easily implemented in circuit analysis programs It applies to any kind of device, such as the bipolar transistor and the MOSFET, whose intrinsic structure operates by charge control Emphasis was given to testing the extrinsic as well as the intrinsic device, and

to the nonidealities present in actual devices. The results of applying this test to BJT devices in a practical inverter circuit are presented.

CHAPTER NINE
THE IMPORTANCE OF THE NON-QUASI-STATIC MODEL
FOR CIRCUIT APPLICATIONS

In this chapter, I address from a practical viewpoint the need for the non-quasi-static models developed earlier. This need depends on the circuit under study. I investigate the usefulness of non-quasi static equivalent circuits in small-signal applications.

9.1 Low-Frequency Applications: Lateral p/n/p Bipolar Transistors

In general, p/n/p transistors are inferior to vertical n/p/n transistors in their current-handling and high-frequency characteristics. However, they are extremely useful in biasing, dc level shifting and serving as active load devices in gain stages for analog integrated circuits [Gro84, Gre84]. Non quasi-static effects can be significant in lateral p/n/p transistors because they have thick quasi-neutral base layers [Ham75], which result in relatively long base transit times.

9.1.1 Case Study: A Series Voltage Regulator

The voltage regulator provides a predetermined dc voltage V_O which is relatively independent of load current, of the temperature, and of variations of the input voltage. In this section, I study the

importance of non-quasi-static effects in lateral p/n/p devices used in a series voltage regulator for automotive applications

The deep base lateral p/n/p devices The lateral p/n/p device is fabricated by a high-voltage bipolar process developed for automotive applications The main characteristics of this high-voltage process is that its n/p/n device has a minimum 40 V V_{CEO} the collector to emitter breakdown voltage at 100 μA collector current The minimum 40 V V_{CEO} is required by the harshest electrical transient, the load dump transient in the automotive environment Load dump [Mac83, Dav73] is a high energy transient caused by an intermittent break in the alternator connector or battery cable This high value of V_{CEO} is achieved by using a fairly thick (16 25 μm) and lightly doped n-epitaxial layer (3 Ω cm) [Gra84]

Figure 9 1(a) shows the layout of minimum geometry (1X) and 4X lateral p/n/p devices In the device layout

1 the emitter is surrounded by the collector so that I maximize the injection from the sidewalls to be collected by the collector
2 the minimum line width is used for the emitter diffusion window (10 μm), to get the largest possible ratio of sidewall to bottom area

The minimum line width also minimizes the undesirable effect of the parasitic emitter-base-substrate p/n/p transistor (as shown in Fig 9 1b), which operates as follows Consider the forward active operation of the lateral p/n/p The parasitic emitter-base substrate transistor is also in the forward-active region and it reduces the effective gain of the lateral device (note that the parasitic collector-base-substrate transistor has both junctions reverse biased)

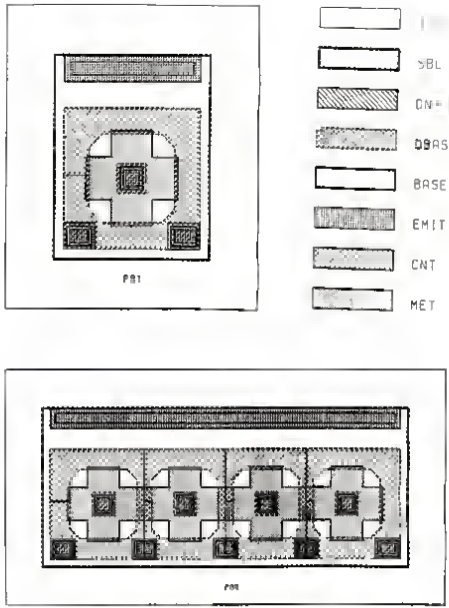
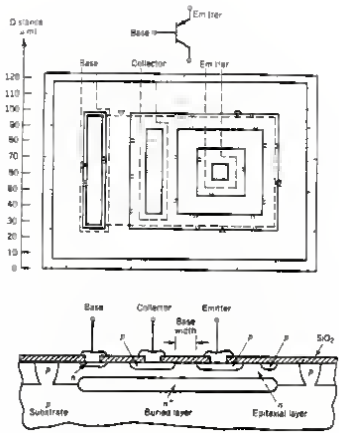


Fig 9.1 Laterel p/n/p devices
(a) The layout of a minimum geometry (1X) and a 4X deep base (D2) laterel p/n/p devices (Not to scale)
(b) Typical device layout and cross section of a laterel p/n/p transistor (after Grebene [Gre84])



(b)

Fig. 9.1 (continued

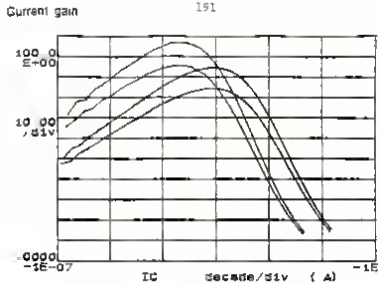
For a regular lateral p/n/p device the p-type emitter and collector are fabricated during the base diffusion cycle for the n/p/n devices. The p/n/p devices in this case study, however, have a deeper base drive for the emitter and collector than that used in regular lateral p/n/p devices. I refer to the devices in Fig. 9.1 as "deep base" (DB) lateral p/n/p devices. Typically, the emitter (and collector) junction depths are 3 to 3.5 μm and 4.5 to 5 μm for regular and DB lateral p/n/p transistors, respectively. The advantages of a DB p/n/p over regular p/n/p devices are

1. collector resistance is about 20 times smaller. As a result, a much smaller $V_{CE(SAT)}$ is achieved,
2. a better current gain is obtained because of the improved emitter injection efficiency, which is contributed by the higher emitter doping density and the larger emitter sidewall area.

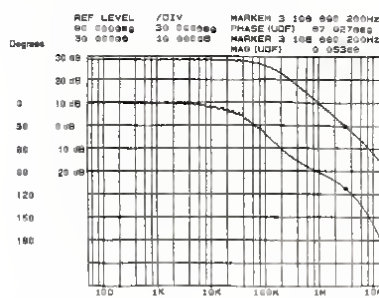
The disadvantages are that the DB structure requires an additional mask and process step and that it uses a larger silicon area (to maintain the same punch-through voltage).

The base width of the DB p/n/p is 18 μm (drawn size) which is the minimum stand-off between the emitter and collector (Fig. 1) for this high voltage process. After the lateral diffusion of the emitter and collector, the base width is approximately 10 μm.

Characteristics of the DB lateral p/n/p devices. Figure 9.2a shows the dc current gain as a function of collector current of the DB p/n/p devices described in Fig. 9.1. The small signal ac gain of the 4X device (Fig. 9.2b) shows a low-frequency current gain (β or h_{21}) of 28 dB and a unity gain frequency (f_T) of 3.1 MHz.



(a)



(b)

Fig 9.2 Characteristics of the DB p/n/p devices (Fig 9.1)
(a) The dc β as a function of collector current. The two curves on the left are for the 1X device, the two curves on the right are for the 4X device. The top and bottom curves for each device are for V_{CB} of 1 and -5 V, respectively
(b) The ac current gain of the 4X device. The top curve is $|\beta|$ and shows a low frequency β of 20 dB and f_T of 3.1 MHz

in Fig 9 3, I compare measured against simulated ac current gain The models used in the simulation are the quasi-static based Gummel Poon model Gummel Poon excess phase model and the small signal L (inductance) model (Fig 4 3C) The first two models are the built in models in SPICE 2G6 The ac parameters are optimized to fit the phase up to 3 f_T (= 10 MHz) The measurement equipment are (1) the HP 3577A Network Analyzer to measure gain characteristic in frequency domain (5 Hz to 200 MHz) (2) the HP 4145B Semiconductor Parameter Analyzer (1 pA to 100 mA, 2 Watt maximum) to determine dc model parameters (3) the Boonton 72BD capacitance meter (0 001 pF to 2000 pF) to measure junction capacitance

As shown in Fig 9 3 the L model accounts for the rapid fall off of $|h_{21}|$ for frequency greater than f_T and is the most accurate model Since the Gummel-Poon model fails to accounts the rapid fall off of $|h_{21}|$, it predicts a lower gain margin (gain at the frequency where the phase shift of h_{21} is 180°)

P/H/P current mirror in a series voltage regulator Figure 9 4a shows a simplified block diagram of a series voltage regulator, which controls the output voltage by controlling the voltage drop across a power transistor the transistor is connected in series between the load and unregulated input voltage The operation of a series regulator is briefly described as follows The internal voltage reference generates a reference voltage, V_R which is independent of the unregulated supply voltage or the temperature variations The error amplifier compares V_R with the sampled and scaled output voltage

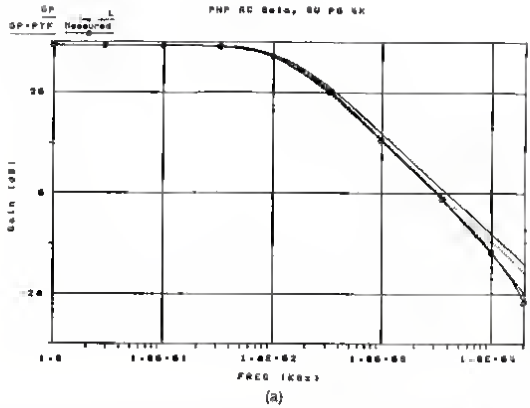
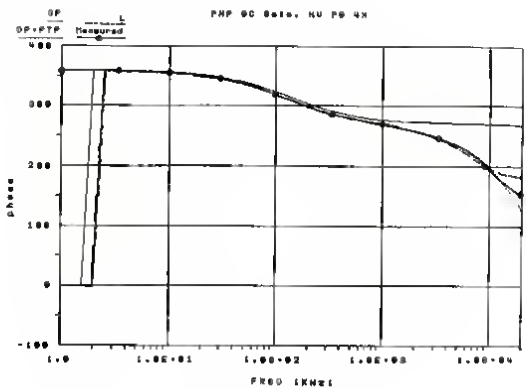


Fig 9 3 The simulated ac current gain against the measurement result GP stands for the small signal Gummel Poon model, GP+PTF is the Gummel Poon model with excess phase and L is the inductance model (Fig 5 3c)
(a) magnitude, (b) phase



(b)

Fig 5 3 continued

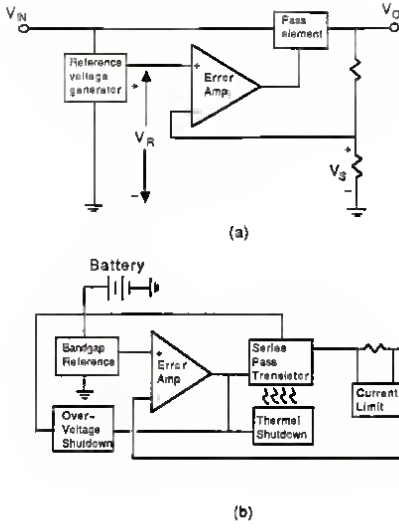


Fig 3 4 A series voltage regulator
(a) A simplified block diagram. (b) A practical series voltage regulator with overload protection circuits for automotive applications

V_S and generates a corrective error signal to regulate the voltage drop across the pass element such that the $V_R \approx V_S$ condition is fulfilled For a detailed analysis see Grebene [Gre84]

A practical voltage regulator with overload protection circuits for automotive applications is shown in Fig 9 4b In this case study the series pass transistor shown in Fig 9.4 is implemented by a current mirror (Fig 9 5) that uses the DE lateral p/n/p devices described earlier The p/n/p device is used as the pass transistor because of the required low dropout voltage (< 1V) between the input and the output, and the possibility of the reverse battery condition If an n/p/n device is used as the transistor the collector is connected to the battery (positive end) and the substrate is grounded (negative end of the battery) In the reverse battery condition, in which the battery is connected in the opposite direction the n type collector is grounded and the p-type substrate is connected to the battery The resulting forward-biased substrate-collector junction generates a very large current that will destroy the circuit On the other hand the emitter of the lateral p/n/p (instead of the collector of an n/p/n) is grounded in a reverse battery condition Consequently, no junction is forward biased for a lateral p/n/p in a reverse battery condition

Note that the p/n/p current mirror is part of the feedback loop, which consists of the pass transistor the output, the voltage divider and the error amplifier (Fig 9 4 and Fig 9 5) Since the p/n/p devices are the slowest devices in the feedback loop their ac gain characteristic determines amount of internal compensation required to ensure the stability of the circuit

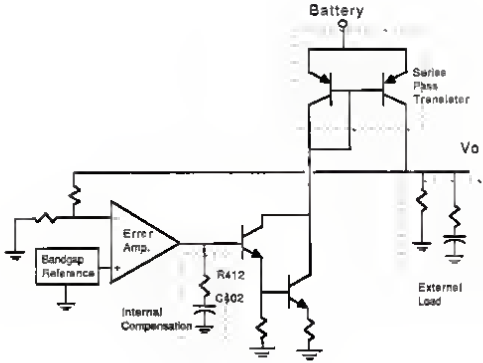
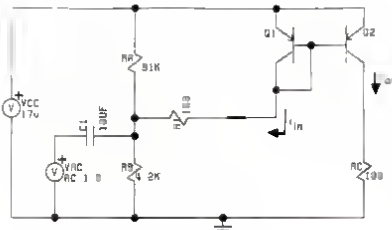


Fig. 9.5 Implementation of the series pass transistor (Fig. 9.4) by a current mirror using the DE lateral p/n/p devices (Fig. 9.1)

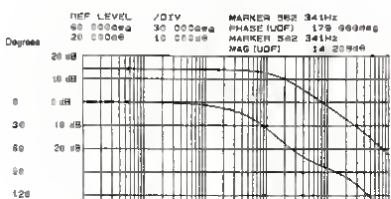
The test circuit shown in Fig 9 6a was built to study the ac characteristics of the p/n/p current mirror. Figure 9 6b shows the ac gain i_{out}/i_{in} of the current mirror (Fig. 9 6a). The ac measurement was performed on an HP 3577A Network Analyzer. The measurement result (Fig. 9 6b) indicates that the low frequency current gain is 14.2 dB, the 3 dB cut off frequency is 170 KHz, and the unity gain frequency is 800 KHz. Note that the low frequency gain is determined by the dc operating point and the emitter area ratio of transistors Q_2 to Q_1 which is, to first order, independent of the gain of p/n/p devices. In contrast, the 3 dB cut off frequency is controlled by the speed of the p/n/p devices. The unity gain frequency is controlled by all the factors just discussed.

Figure 9 7 shows a computer simulation of the ac gain of the p/n/p current mirror. Again, the small-signal Gummel-Poon, the Gummel-Poon with excess phase, and the L models are used in the simulation. As in the simulation for the DS lateral p/n/p devices (Fig. 9 3), the L model best accounts for the rapid fall-off of the current gain beyond the unity gain frequency and is the most accurate model.

Internal compensation of a series voltage regulator. Figure 9 8 shows the ac gain of the feedback loop in the series voltage regulator (Fig. 9 5). The external load contributes the low-frequency pole (33 Hz) and zero (3.3 KHz). The ac loop gain before compensation, shown in Fig. 9 8a, indicates that there is not enough phase margin (= 25°) at 5 MHz. Therefore the internal compensation circuit, which consists of a capacitor, C402, and a resistor, R412, as indicated in Fig. 9 5, is added to provide a pole and a zero to ensure a minimum 45 degree phase

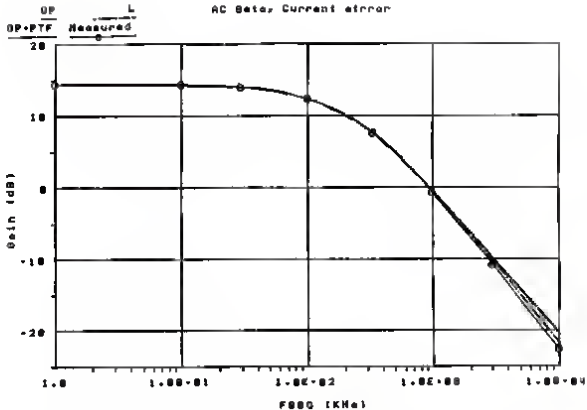


(a)



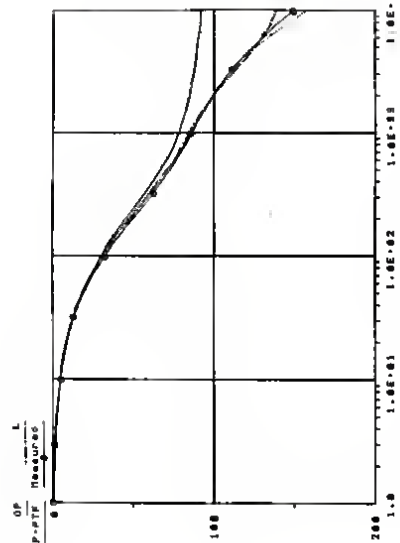
(b)

Fig 9.6 Characteristics of the p/n/p current mirror.
(a) The test circuit used for ac characterization.
(b) The measured ac current gain defined by i_{out}/i_{in}. The top curve is the magnitude and the bottom curve is the phase



(a)

Fig 3.7 The simulated ac β against the measured data for the p/n/p current mirror. GP stands for the small signal Gummel Poon model. GP+PTF is the Gummel-Poon model with excess phase, and L is the inductance model (a) Magnitude; (b) Phase



(b)

Fig. 9.7- continued

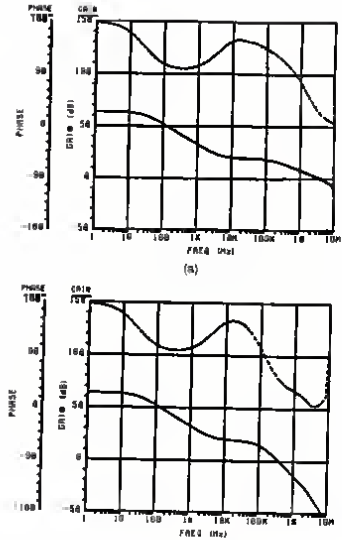


Fig 9 8 The ac gain simulated using SPICE Gummel Poon model with excess phase of the feedback loop in the series voltage regulator (Fig 9 5) The top and bottom curves are the phase and the magnitude of the gain respectively (a) Before compensation (b) After compensation

margin The on-chip compensation pole (≈ 100 KHz) is contributed by the 75 pF C402 in combination with the output impedance of the error amplifier in parallel with the input impedance of Darlington stage which is about 20 KΩ The compensation zero (≈ 1 MHz) is contributed by C402 and R412 (2 KΩ)

The simulated ac loop gain (Fig 9 8) uses the SPICE 2G6 built-in Gummel-Poon model with excess phase As discussed earlier this model fails to predict the rapid fall-off of the magnitude of the current gain for frequency greater than 3 MHz Since the rapid fall-off reduces the amount of internal compensation required, the quasi-static based Gummel Poon model leads us to overdesign more internal compensation is provided than is needed Excess compensation wastes silicon area and reduces the useful operating frequency range

The use of the L model which accurately predicts the rapid fall-off of the magnitude of the current gain, predicts that a 200 KHz pole instead of a 100 KHz one is required Use of the L model thus reduces by a factor of two the silicon area used for the relevant capacitor Since a 75 pF junction capacitor requires silicon area of $3 \cdot 10^5\ \mu m^2$ the saving is $1\ 5\ 10^5\ \mu m^2$ which is about the size of five DE lateral p/n/p devices

5 1 2 Level Shifting in an Operational Amplifier

A lateral p/n/p transistor in the common base configuration used for dc level shifting in a high voltage operational amplifier is investigated here The level shifting circuit is shown in Fig 9 9

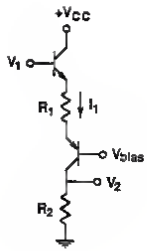


Fig 9 9 A commonly used p/n/p-n/p/n level-shift stage. The lateral p/n/p transistor is operated in common-base configuration to enhance its frequency capability.

The lateral p/n/p device has a base width of 22.5 μm in which a high-resistivity epitaxial layer is used to obtain a good current gain (≈ 100). The cutoff frequency (f_α) is $2 \cdot 10^6$ Hz. The lateral p/n/p device is in the signal path of the operational amplifier [Gra84]. Since in the design of a feedback amplifier it must be ascertained that the circuit is stable at all frequencies including $f > f_T$ and not merely over the frequency range of interest [Mil86], the lateral p/n/p device is important to the stability of the overall amplifier.

Figure 9.10 shows the measured phase and magnitude of h_{21} of the lateral p/n/p device, and the results of simulation using the SPICE 2G6 built-in model (the Gummel-Poon model), the Gummel-Poon model with excess phase [Tho88], and the small-signal L model (Fig. 4.3c). The base transit time used in each model is optimized to closely match the measured phase up to 10 MHz.

Comparing the hybrid-π model including excess phase with the L model, I note that at 10 MHz, where the measured phase shift is 180°, the hybrid-π model predicts $|h_{21}|$ of -15 dB, and the L model predicts $|h_{21}|$ of -21 dB. The measured $|h_{21}|$ is -20 dB. The operational amplifier that employs this lateral p/n/p device is stable because the steeply falling $|h_{21}|$ above f_T provides the necessary gain margin (Fig. 9.10b). Since both of the hybrid-π models fail to simulate this steeply falling $|h_{21}|$, they predict instability of the operational amplifier, which is incorrect. Only the L model accurately predicts both the phase and the magnitude of h_{21}.

This case study is due to T. Ifstrøm, who used our L model to get an accurate computer simulation of the operational amplifier. The

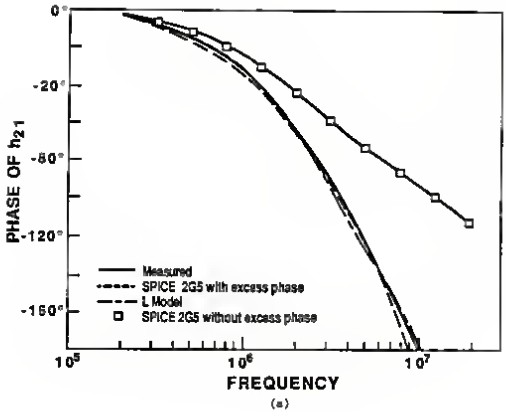
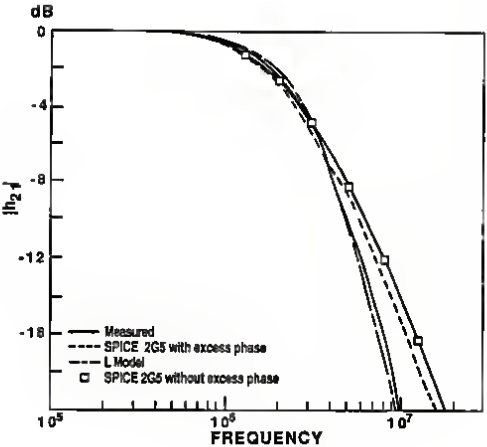


Fig 9.10 The measured and simulated h_{21} of the lateral p/n/p transistor in Fig 9.9
(a) Phase (b) Magnitude



(b)

Fig. 9 10--continued

results are consistent with the discussion about self-consistency associated with (§ 11)

§ 2 High-Frequency Small Signal Application

As discussed in Chapter Two, at high frequencies the quasi-static approximation becomes invalid because the carriers in the base region have insufficient time to redistribute. As a result, the measured small-signal transistor behavior is poorly predicted by the hybrid π model at high frequencies [Gou82]. I shall overview approaches to high-frequency (including microwave) transistor modeling and simulation to investigate the need for small-signal non-quasi-static models for high frequency applications.

Macnee and Telsky [Mac72] compared the measured high frequency values of y_{11} and y_{21} for the 2N918 transistor (f_T = 900MHz) with the simulation results predicted by the hybrid-pi model with excess phase. They found that a discrepancy persisted in both y_{11} and y_{21} after adding lead capacitance and inductance to the hybrid pi model with excess phase. They then modified the hybrid π model as shown in Fig. § 11, which used a multi-section model for the quasi-neutral base region. The model parameters were determined by optimizing y_{11} and y_{21} up to $f_T/2$. The resulting model (Fig. § 11) agreed more closely with the measured admittances.

Ollins and Ratner [Oll72] derived the e-parameters for the hybrid π model (including excess phase), in which the element values were optimized to attain a good agreement with experimental measurement data

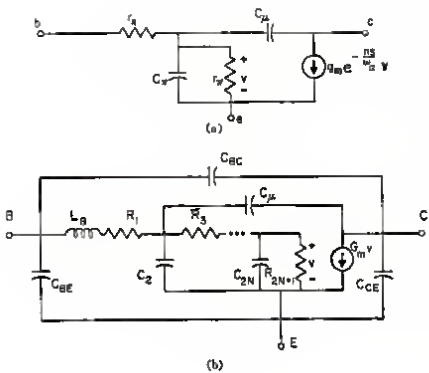


Fig 9.11 Two models used in high frequency applications After Macnee and Talsky [Mac72]
(a) The hybrid-π model (b) N-lump extension plus extrinsic elements

over a frequency range from 100 MHz to 1 GHz. In comparing the results with the experimental data the major differences were in the magnitude of s_{11} and the phase of s_{21}. They concluded that to obtain a closer match between the model and the measured s parameters a distributed model for the base region was needed. The resulting distributed model was similar to that of Macnee and Teleky.

Gough [Gou82] measured the s-parameters of the HP35826E microwave transistor with f_T = 4 GHz. He then compared the experimental data with that predicted by the hybrid π model (no excess phase) including extrinsic elements such as lead inductance and capacitance. (All the parameters except r_e and β_0 were optimized.) Gough then improved the hybrid-π model by adding a one-section R-C ladder equivalent circuit for the base layer.

Schutt-Aine [Sch83] derived the short circuit input impedance of an ion implanted microwave transistor using its geometrical properties. He developed both a hybrid-π model and a distributed model. The hybrid π equivalent circuit was found to agree up to 0.5 GHz with the measured data. A forty section distributed model (Fig. 9.12) for the base layer was used and good agreement was maintained up to 10 GHz.

From the previous work on the high frequency transistor modelling, we see that the quasi-static model (the hybrid π model) is invalid at high frequencies and that multi-section models for the base layer are commonly employed (see also Kakihana and Wang [Kak71] and Lange and Carr [Lan72]). Using the exact solution as the standard for comparison, we have shown in Chapter Four that non-quasi-static multisection models require half as many sections (lumps) as

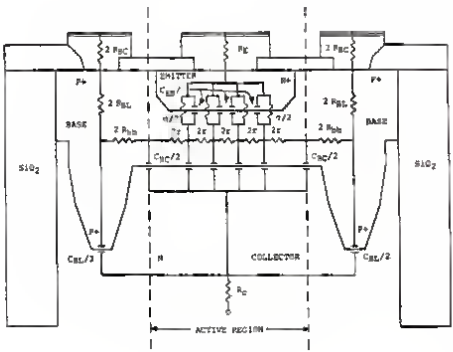


Fig. 9.12 Cross section of double-base strip single-emitter strip microwave transistor showing the distributed intrinsic parameters and extrinsic elements. After Schutt Ainé [Sch83]

quasi static multisection models To accurately approximate the exact solution using only a few circuit elements or centralized sources which is a primary goal of circuit simulation modeling I recommend using the non-quasi-static multisection models (see Chapter Four) instead of the commonly used distributed quasi-static model

§ 3 Summary

I investigated lateral p/n/p transistors used as a series pass transistor in a series voltage regulator and as a dc level-shifter in a high-voltage operational amplifier In the case study of the series voltage regulator, the non-quasi static L model correctly predicts the rapid fall-off of the magnitude of the current gain and thus reduces the required internal compensation For the high voltage operational amplifier, only the non-quasi-static L model correctly predicts the observed stability I then review the microwave literature for indications of non quasi static effects that other workers have claimed to be important in practical circuits Such indications exist in the microwave transistor literature even though the base layers involved have a thickness less than 1000 Å [Sch83]

CHAPTER TEN
SUMMARY AND RECOMMENDATIONS

The major accomplishments of this study are the

1) Review and unification of previous work related to the use of the quasi-static approximation in modeling the quasi-neutral layers of p/n junction devices

2) Improvement of analytical methods for determining delay time. The area and ramp approaches, which are new, are developed for determining the Elmore delay time. Taking into account the delay associated with the convection current and the storage current determined by the newly developed methods, I improved the small-signal quasi static model for a junction diode

3) Modification of existing small signal quasi-static models for diodes and transistors. The resulting new quasi-static models represented the two-port parameters, especially y_{11} and y_{21}, more accurately than had previous models

4) Development of a systematic approach to include propagation delay in quasi-neutral layers. Despite the detailed physics included, the resulting large signal models are compact enough for use in circuit simulators

5) Comparison of recent one dimensional bipolar transistor models. I showed that the TIGC, PCM and NC models did not give the correct trend for $|y_{21}|$. The L model is more accurate, yet no more complex

6) Implementation of the new large-signal non quasi-static models (the

L and NC models) in a circuit simulation program using the user defined control approach. The non-quasi-static models accurately described the delay in digital circuit analysis. The non quasi static models established in a circuit simulator are used to assess the importance of non-quasi static models in digital applications.

7) Proposal of a test for the self-consistent validity of the quasi static approximation. It demonstrated the usefulness of the self-consistency test for a power bipolar transistor.

8) Investigation of lateral p/n/p transistors in a series voltage regulator and in operational amplifiers. The non quasi-static models for lateral p/n/p transistor accurately predicted the gain margin and stability of the operational amplifiers, while previously available models did not.

I suggest the following studies based on the methods and approaches of the present study.

1) Develop experimental techniques to accurately extract the delay time associated with the convection current.

2) Assess (experimentally) the determination of parameters that are based on the quasi-static model such as the input circuit method [Get78] to determine, r_b, the extrinsic base resistance.

3) Write the non-quasi static models into circuit simulation programs to provide a better simulation efficiency.

APPENDIX A
MATHEMATICAL DETAILS OF DERIVATION OF
THE LARGE SIGNAL L MODEL

In this appendix I provide the mathematical details in the derivation of the large signal L model in Section 5.4.2. We shall first explain how I obtain Eqn. (5.36). I then derive (5.43) from (5.42). I complete this appendix by describing the mathematical manipulation from (5.43) to (5.38).

Solution of (5.36). From a theorem in [Ras74] the ordinary differential equation shown in (A.1) has a solution shown in (A.2):

$$\frac{di(t)}{dt} + P(t)\, i(t) = Q(t) \tag{A.1}$$

$$i(t) = e^{\int P(t)\, dt} \left[\int e^{\int P(t)\, dt}\, Q(t)\, dt + C \right] \tag{A.2}$$

Since $P(t) = R'/L' = 1/\tau_D$ and $Q(t) = n_E(t)/L'$, the solution for $i(t)$ in Eq. (5.35) is given in (A.3):

$$i(t) = e^{-t/\tau_D} \left[\int e^{t/\tau_D} \frac{n_E(t)}{L'}\, dt + C \right] \tag{A.3}$$

The constant C is obtained by letting $i(t=0) = i(0)$:

$$C = i(0) - \left[\int e^{t/\tau_D} \frac{n_E(t)}{L'}\, dt \right]_{t=0}$$

$$= i(0) - F(0) \tag{A.4}$$

where F(t) is defined by Eq (5 37) Substituting [illegible] yields Eq (5 36)

Equivalence between (5.42) and (5.43) To prove (5 42) is equivalent to (5 43), I need to prove the equivalence shown in (A 5) is true

$$\frac{e^{-t/r_D}}{R'_{QS}} \int e^{\frac{t}{r_D}} \frac{dn_g(t)}{dt}\, dt = \frac{d}{dt}\left[\frac{e^{-t/r_D}}{R'_{QS}} \int e^{\frac{t}{r_D}} n_g(t)\, dt\right] \qquad (A\ 5)$$

The proof starts with rewriting the right-hand-side of (A 5) by differentiation

$$\frac{d}{dt}\left[\frac{e^{-t/r_D}}{R'_{QS}} \int e^{\frac{t}{r_D}} n_g(t)\, dt\right] = \frac{1}{R'_{QS}}\left[\frac{e^{t/r_D}}{r_D} \int e^{\frac{t}{r_D}} n_g(t)\, dt + n_g(t)\right] \qquad (A\ 6)$$

Next I shall rewrite the left-hand-side of (A 5) as (A 8), which needs to utilize (A 7)

$$\int \frac{d}{dt}\left[e^{t/r_D} n_g\right] dt = \int \left(\frac{e^{t/r_D}}{r_D} n_g + e^{t/r_D} \frac{dn_g}{dt}\right) dt \qquad (A\ 7)$$

$$\frac{e^{-t/r_D}}{R'_{QS}} \int e^{\frac{t}{r_D}} \frac{dn_g(t)}{dt}\, dt = \frac{e^{-t/r_D}}{R'_{QS}}\left[\int \frac{d}{dt}\left[e^{\frac{t}{r_D}} n_g(t)\right] dt - \int \frac{e^{\frac{t}{r_D}}}{r_D} n_g\, dt\right]$$

$$= \frac{n_g}{R'_{QS}} - \frac{e^{-t/r_D}}{R'_{QS}\, r_D} \int e^{\frac{t}{r_D}} n_g(t)\, dt \qquad (A\ 8)$$

Since (A 8) = (A 6), I complete the proof that (5 42) is equivalent to (5 43)

Transform indefinite integral into definite integral In this section we shall derive Eq (5 38) from Eq (5 43) Let $G(t) = \int g(t)\,dt$ then by definition

$$\int_0^t g(\eta)\,d\eta = G(t) - G(0) \tag{A 9}$$

Let $g(t) = \exp(t/\tau_D)\,n_E$ then

$$\int g(t)\,dt = \int_0^t e^{\eta/\tau_D} n_E(\eta)\,d\eta + G(0) \tag{A 10}$$

where

$$G(0) = \left[\int e^{t/\tau_D} n_E(t)\,dt\right]_{t=0} = F(0)/L',$$

in which F(0) is defined in (5 37) Substituting (A 10) into (5 43) yields

$$i(t) = \frac{n_E(t)}{\tau_{QS}} - \frac{d}{dt}\left[e^{-t/\tau_D}\left[\frac{1}{\tau_{QS}}\int_0^t e^{\frac{\eta}{\tau_D}} n_E(\eta)\,d\eta + C(0)\right]\right] + e^{-\frac{t}{\tau_D}}\left[i(0) - \frac{G(0)}{L'}\right], \tag{A 11}$$

Note that the the terms involving G(0) in (A 11) cancel Combining the last two terms in (A 11) yields (A 12)

$$i(t) = \frac{n_E(t)}{\tau_{QS}} - \frac{d}{dt}\left[\frac{e^{t/\tau_D}}{\tau_{QS}}\int_0^t e^{\frac{\eta}{\tau_D}} n_E(\eta)\,d\eta - e^{\frac{t}{\tau_D}} \tau_D\, i(0)\right] \tag{A 12}$$

Since Eq (A 12) = Eq (5 38), I complete the derivation from (5 43) to (5 38)

APPENDIX B
UDCS/SLICE/SPICE MODEL IMPLEMENTATION

In this appendix we give the FORTRAN code listing of the implementation of $dq_{C(NC)}/dt$ and $dq_{C(L)}/dt$. This serves as a brief overview of a representative UDCS/SLICE/SPICE implementation of a charge-based network representation for the large-signal NC and L models. Readers are refered to Yuan and Elsenstadt [Yua87] for other current sources in the NC or L models.

The theory and mechanical details behind the implementation of steady state network models into SLICE/SPICE can be found in the SLICE users manual [Her86]. The theory behind the implementation of transient UDCS network models in SLICE is discussed in detail in Veeraraghvan and coworkers [Vee86].

FORTRAN code listing for the implementation of $dq_{C(NC)}/dt$.

```
C
C
C  The following listing is a subroutine which calculates the
C  charging current accounted for the delay associated with the
C  non quasi-static effects in the large signal NC model
C
C
      SUBROUTINE UUCTS(Y IFLAG LF NP LC NC JERROR)
      IMPLICIT DOUBLE PRECISION (A-H O-Z)
      SPECIAL COMMON BLANK
      COMMON /KNSTNT/ TWOPI XLOG2,XLOG10 ROOT2 RAD BOLTZ CHARGE CTOK
     1   GMIN RELTOL ABSTOL VNTOL TRTOL CHGTOL EPSO EPSSIL EPSOX
      COMMON /STATUS/ OMEGA TIME DELTA OLDLD(7) AC(7) VT XNI EGFET
     1  MODE MODEDC ICALC INITF METHOD IORD,MAXORD NONCON ITERNO
     2  ITEMNO NOSOLV
      COMMON /BLANK/ X(64)
C
C
C
C     The following is the format used for circuit files  Consult
C     the SLICE users manual for details
C
C     Gxx N2 N1 POLY 1 NB NE EXTERN UUCTS DVBE DVBC QN QN 1 IN QTEMP
C        ITEMP DELTEMP DELPR AREA CJEO ME PBE VBO VA VAL DOPE DN
C
C
      GOTO (100 200 300)  IFLAG+2
100   IF (NC NE 1) JERRDR =  99010
      RETURN
200   VBE=X(LC+1)
      AREA=X(LP+10)
      VBO=X(LP+14)
      DOPE=X(LP+17)
      DN=X(LP+18)
      TAUD=VBO*VBO/6 0/DN
C
C     q n_i^2 = 33 64 & kT/q = 0 2585
C
      Q=AREA*33 64*DN/DOPE/VBO*TAUD*(DEXP(VBE/D 02585)-1 0)
      DL=AREA*33 64*DN/DOPE/VBO*TAUD*DEXP(VBE/D 02585)/0 02585
      IF (TIME EQ 0 0) THEN
         X(LP+3)=Q
         X(LP+4)=Q
         X(LP+5)=0 0
         X(LP+6)=Q
         X(LP+7)=0 0
         X(LP+8)=0 0
         X(LP+9)=0 0
```

```
        Y=0.0
        RETURN
      ENDIF
C - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
      IF (INITF.EQ.6 AND.DELTA.GE.X(LP+8)) THEN
        X(LP+4)=X(LP+3)
        X(LP+3)=X(LP+6)
        X(LP+5)=X(LP+7)
        X(LP+9)=X(LP+8)
      ENDIF
C - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
      IF ((METHOD.EQ.1).OR.(METHOD.EQ.2.AND.X(LP+9).EQ.0.0)) THEN
        Y=2.0/DELTA*(Q-X(LP+3))-X(LP+5)
        X(LP+1)=2.0*D1/DELTA
      ELSE
        F1=(2.0+X(LP+9)/DELTA)/(X(LP+9)+DELTA)
        F2=DELTA/X(LP+9)/(X(LP+9)+DELTA)
        Y=F1*Q-(F1+F2)*X(LP+3)+F2*X(LP+4)
        X(LP+1)=F1*D1
      ENDIF
C - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
      X(LP+6)=Q
      X(LP+7)=Y
      X(LP+8)=DELTA
      RETURN
300   Y=X(LP+1)
      RETURN
      END
```

FORTRAN code listing for the implementation of $dq_{C(L)}/dt$.

```
C-----------------------------------------------------------------
C  The following listing is a subroutine which calculates the
C  charging current accounted for the delay associated with the
C  non-quasi-static effects for the large signal L model
C-----------------------------------------------------------------
C
      SUBROUTINE UUQL(Y,IFLAG,LF,NE,LC,NC,IERROR)
      IMPLICIT DOUBLE PRECISION (A-H,O-Z)
      SPECIAL COMMON /BLANK/
      COMMON /KNSTNT/ TWOPI,XLOG2,XLOG10,ROOT2,RAD,BOLTZ,CHARGE,CTOK,
     1    CMIN,RELTOL,ABSTOL,VNTOL,TRTOL,CHGTOL,EPSO,EPSSIL,EPSOX
      COMMON /STATUS/ OMEGA,TIME,DELTA,DELOLD(7),AG(7),VT,XNI,EGFET,
     1  MODE,MODEDC,ICALC,INITF,METHOD,LORD,MAXORD,NONCON,ITERNO,
     2  ITEMNO,NOSOLV
      COMMON /$LANK/ X(64)
C
C
C
C     The following is the format used for circuit files.  Consult
C     the SLICE users manual for details
C
C     GXXX N2 N1 POLY 1 N8 NE EXTERN UUQL DVBE DVBC QN QN 1 IN QTEMP
C       1TEMP DELTEMP DELTR AREA CJEO ME FBE VBO VA VAI DOPE DN
C       VBEN 1 VN 1EMP TRITN 1 TRITEMP DTRIN 1 DTRTEMP
C-----------------------------------------------------------------
C
      GOTO (100,200,300),IFLAG+2
100   IF (NC.NE.1) IERROR = 39010
      RETURN
200   VBE=X(LC+1)
      VBO=X(LF+14)
      DN=X(LF+18)
      AREA=X(LF+10)
      DOPE=X(LF+17)
      TAUD=VBO*VBO/6.0/DN
C-----------------------------------------------------------------
C     q n_i^2 = 33.64 & kT/q = 0.2585
C-----------------------------------------------------------------
      IF (TIME.EQ.0.0) THEN
         CIZERO=33.64*AREA/DOPE*DN/VBO*
     1    (DEXP(VBE/0.02585)-1.0)
         Q=CIZERO*TAUD
         X(LF+3)=Q
         X(LF+4)=-Q
         X(LF+5)=0.0
         X(LF+6)=-Q
         X(LF+7)=0.0
         X(LF+8)=0.0
```

```
      X(LP+8)=D.0
      X(LP+19)=V8E
      X(LP+2D)=V5E
      X(LP+21)=0.0
      X(LP+22)=D.D
      X(LP+23)=D.D
      X(LP+24)=0.D
      Y=D.D
      RETURN
    ENDIF
CCC---------------------------------
    IF (1N1TF EQ 6 AND DELTA GE X(LP+8)) THEN
      X(LP+4)=X(LP+3)
      X(LP+3)=X(LP+6)
      X(LP+5)=X(LP+7)
      X(LP+9)=X(LP+8)
      X(LP+19)=X(LP+2D)
      X(LP+21)=X(LP+22)
      X(LP+23)=X(LP+24)
    ENDIF
    V8ELST=X(LP+19)
    TRINT=X(LP+21)*DEXP(-DELTA/TAUD)+DELTA/2.0*(
   1      DEXP(-DELTA/TAUD)*(DEXP(V8ELST/D.02585)-1.0)+
   2      (DEXP(V8E/D.02585)-1.D))
    DTRINT=X(LP+23)*DEXP(-DELTA/TAUD)+DELTA/2.D*(
   1       DEXP(-DELTA/TAUD)*DEXP(V8ELST/0.02585)+
   2       DEXP(V8E/0.D2585))
    Q=AREA*33.64/DOPE/WBO*DN*TRINT+
   1  C1ZERO*TAUD*DEXP(-TIME/TAUD)
    D1=AREA*33.64/DOPE/WBO*DN/0.02585*DTRINT
CCC---------------------------------
    IF ((METHOD EQ 1) OR (METHOD EQ 2 AND X(LP+9) EQ D.D)) THEN
      Y=2.D/DELTA*(Q-X(LP+3))-X(LP+5)
      X(LP+1)=2.D*D1/D8LTA
    ELSE
      F1=(2.0+X(LP+9)/DELTA)/(X(LP+9)+DELTA)
      F2=DELTA/X(LP+9)/(X(LP+9)+DELTA)
      Y=F1*Q-(F1+F2)*X(LP+3)+F2*X(LP+4)
      X(LP+1)=F1*D1
    ENDIF
CCC---------------------------------
    X(LP+6)=Q
    X(LP+7)=Y
    X(LP+8)=DELTA
    X(LP+2D)=V8E
    X(LP+22)=TRINT
    X(LP+24)=DTRINT
    RETURN
3DD Y=X(LP+1)
    RETURN
    END
```

APPENDIX C
TRANSFORMATION OF THE NC AND L MODELS TO
CHARGE CONTROL AND PCE REPRESENTATIONS

To transform the large-signal NC and L models developed in Appendix A to the corresponding charge-control and PCE representations we start by redrawing the negative capacitance model (Fig 5 4) and the L model (Fig 5 5) in terms of controlled sources (Fig C 1a)

NC model The values of the elements in Fig C 1(a) are

$$|i_{QS}| = v_{BE}/R_{QS} \quad , \tag{C 1}$$

which is given in (5 29)

$$|q_{C(NC)}| = \left|\int_0^{v_{BE}} -C\, dv\right| = |i_{QS}|\, \tau_D \quad , \tag{C 2}$$

and

$$|q_{QS}| = \int_0^{v_{BE}} C_{QS}\, dv = |i_{QS}|\, \tau_{TF} \tag{C 3}$$

We first replace the dependent current source i in Fig C 1(a) by $|i_{QS}|$ and dq_C/dt. We then combine the two charging currents dq_{QS}/dt and dq_C/dt into one charging current dq_E/dt. The result is Fig C 1(b), where

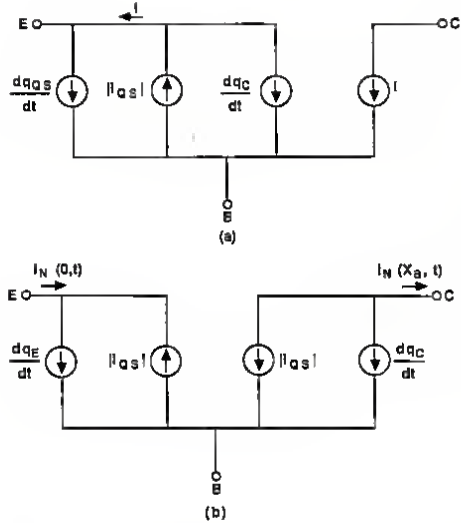


Fig C 1 The negative capacitance and L models in terms of controlled sources
(a) Transformation of Fig. 5 3 to controlled sources representation,
(b) The equivalent circuit in Fig C 1(a) transformed by first replacing i by $|i_{QS}|$ and dq_C/dt and then combining dq_{QS}/dt and dq_C/dt into dq_E/dt

$$q_{E(NC)} = q_{QS} - q_{C(NC)} - i_{QS}\frac{X_B^2}{D_n}\left(\frac{e^{\eta} - e^{-\eta} + 2\eta}{\eta^2(e^{\eta} - e^{-\eta} - 1)}\right) \qquad \text{(C 4)}$$

In Fig C-1(b), the two conduction current sources make no contribution to i_B. Thus we can redraw Fig. C-1(b) as Fig. 6-2, which is the FCB configuration. For a uniformly doped base layer, i_{QS} is defined in (5-33)

$$q_{C(NC)} = i_{QS} X_B^2/(5 D_n) = q_{QS}/3, \qquad \text{(C-5)}$$

and

$$q_{E(NC)} = i_{QS} X_B^2/(3 D_n) = 2q_{QS}/3 \qquad \text{(C-6)}$$

where $q_{QS} = A_E q n_B X_B/2$ for $\eta = 0$.

L model. Equations (5-28) and (5-40) together with (C-3) and

$$q_{E(L)} = q_{QS} - q_{C(L)} \qquad \text{(C-7)}$$

correspond to an L model in the same FCB form as shown in Fig. 6-2.

APPENDIX D
DETERMINATION OF ELEMENTS OF TICC AND PCB MODELS
FOR AN EXPONENTIALLY DOPED BASE

We shall derive i_{QS}, q_{QS}, q_E and q_C for the TICC and PCB models for an exponentially doped base layer, that is,

$$N_A(x) = N_A(0)\exp(-\eta x/X_B) \tag{D 1}$$

in low level injection. For this derivation we need the quasi-static excess electron density $n_{QS}(x,t)$ in the quasi-neutral base layer. This is obtained by solving the electron continuity equation in steady state and then applying the quasi static approximation

$$n_{QS}(x,t) = \frac{|i_{QS}|\, X_B}{q\, A\, D_n} \frac{1 - e^{-\eta(1 - x/X_B)}}{\eta} \tag{D 2}$$

where $|i_{QS}|$, given in (5.29), is position independent because of the assumed absence of recombination.

TICC. The charges q_E and q_C for the TICC model are given in (4) of Klose and Wieder [Klo87]

$$q_{E(TICC)} = -q\, A \int_0^{X_B} FE(x)\, n(x,t)\,dx \tag{D 3}$$

and

$$q_{C(TICC)} = q A \int_0^{X_E} FC(x)\, n(x,t)\, dx \qquad (D\ 4)$$

To determine FE(x) and FC(x) requires the knowledge of the hole distribution in both in space and time, which can be assumed to be (D 1) for low injection. Substituting (D 1) into (A13d) and (A13e) in Klose and Wieder [Klo87] we obtain FE(x) and FC(x)

$$FE(x) = \frac{e^{\eta x/X_B} - e^{-\eta}}{1 - e^{-\eta}} \qquad (D\ 5)$$

and

$$FC(x) = \frac{1 - e^{-\eta x/X_B}}{1 - e^{-\eta}} \qquad (D\ 6)$$

The TICC model must make use of the quasi static approximation to obtain analytical expressions for q_E and q_C. $n(x,t) \rightarrow n_{QS}(x,t)$ Substituting (D 2) and (D 6) into (D 4) yields

$$q_{C(TICC)} = i_{QS} \frac{X_B^2}{D_n} \left(\frac{1 + e^{-\eta}}{\eta\,(1 - e^{-\eta})} - \frac{2}{\eta^2} \right) \qquad (D\ 7)$$

Similarly, substituting (D 2) and (D 5) into (D 3) yields

$$q_{E(TICC)} = i_{QS} \frac{X_B^2}{D_n} \left(\frac{e^{\eta} - e^{-\eta} + 2\eta}{\eta^2 (e^{-\eta} - 1)} \right) \qquad (D\ 8)$$

Note that $q_{C(TICC)} + q_{E(TICC)} = q_{QS}$ which is consistent with (B 3)

The conduction current for TICC model is given by the Gummel formulation [Gum70a]

$$|i_{QS}| = \frac{A\, q\, D_n\, n_i^2}{\int_0^{X_B} N_A(x)\, dx} \left[\exp\left(\frac{V_{BE}(c)}{kT/q}\right) - 1\right]$$

$$= \frac{A\, q\, D_n\, n_i^2\, \eta}{N_A(0)\, X_B\, (1 - e^{-\eta})} \left[\exp\left(\frac{V_{BE}(c)}{kT/q}\right) - 1\right] \qquad (D\ 9)$$

which is equivalent to (5.29)

FCS Fossum and Veeraraghavan [Fos86] defined the base charge components for the FCS model

$$q_{E(FCS)} = -q\, A \int_0^{X_B} \left(1 - \frac{x}{X_B}\right) n_{QS}(x, t)\, dx \qquad (D\ 10)$$

and

$$q_{G(FCS)} = -q\, A \int_0^{X_B} \left(\frac{x}{X_B}\right) n_{QS}(x, t)\, dx \qquad (D\ 11)$$

Substituting (D.2) into (D.11) yields

$$q_{G(FCS)} = i_{QS} \frac{X_B^2}{D_n} \left[\frac{1}{2\eta} - \frac{1}{\eta^2} + \frac{1}{\eta^3}(1 - e^{-\eta})\right] \qquad (D.12)$$

Similarly, substituting (D.2) into (D.10) yields

$$q_{E(FCS)} = i_{QS} \frac{X_B^2}{D_n} \left[\frac{\eta^2 - 2 + 2\, e^{-\eta}(1+\eta)}{2\, \eta^3}\right] \qquad (D.13)$$

Equation (D 12) and (D 13) are equivalent to (D 2) and (D 3) only for the special case of a uniformly doped base layer The conduction current for the PCB model, which is the space averaged electron current, is

$$\frac{1}{X_B}\int_0^{X_B} i_N(x,t)\,dx = \frac{q\,D_n}{X_B}\left[n_E - \frac{\eta}{X_B}\int_0^{X_B} n(x\ t)\,dx\right] \qquad \text{(D 14)}$$

To evaluate (D 14) requires the knowledge of n(x,t) during transients Thus (D 14) is equivalent to (5 29) and (D 9) only for a uniformly doped base layer ($\eta = 0$).

REFERENCES

Ant86 Antognetti, P., Power Integrated Circuits, Physics, Design, and Applications, pp 247-252, McGraw-Hill, New York, 1986

Arr77 Arreole, J. L., and F. A. Lindholm, "A Nonlinear Indefinite Admittance Matrix for Modeling Electronic Devices," IEEE Trans. on Electron Devices, vol. ED-24, pp. 756-767, June 1977

Ash64 Ashar, K. G., "The Method of Estimating Delay in Switching Circuits and the Figure of Merit of a Switching Transistor," IEEE Trans. on Electron Devices, vol. ED-11, pp. 497-506, Nov 1964

Bae57 Beaufoy, R., and J. J. Sparkes, "The Junction Transistor as a Charge Controlled Device," ATE J., vol. 13, pp. 310-324, Feb 1957

Bul68 Bulucea, C. D., "Diffusion Capacitance of p-n Junctions and Transistors," Electronic Letters, vol. 4, no. 25, pp. 559-561, 13th Dec. 1968.

Che86 Chen, M. K., "Inclusion of Quasineutral-Region Delay in Quasi-Static Diode Analysis," Master's Thesis, University of Florida, 1986

Che87 Chen, M. K., F. A. Lindholm, and T. W. Jung, "Non Quasi-Static Small-Signal Models for Semiconductor Junction Diodes with Extensions for Transistors," Solid-State Electronics, vol. 30, no. 8, pp. 883-885, 1987.

Che88 Chen, M. K., F. A. Lindholm, and S. S. Wu, "Comparison and Extension of Recent One-Dimensional Transistor Models," IEEE Trans. on Electron Devices, vol. ED-35, pp. 1096-1106, June 1988.

Che71 Cherry, E. M., "Active-Device Capacitance," IEEE Trans. on Electron Devices, vol. ED-18, pp. 1166-1168, Dec. 1971

Chu75 Chua, L. O., and P. M. Lin, Computer Aided Analysis of Electronic Circuits, Algorithms & Computational Techniques, Prentice Hall, Inc., Englewood Cliffs, New Jersey, 1975

Dav73 Davis W F , "Bipolar Design Considerations for the Automotive Environment " IEEE J. Solid-State Circuits, vol SC-8 pp 419 426 Dec 1973

Deg85 De Greff H C , and W J Kleesterman "New Formulation of the Current and Charge Relations in Bipolar Transistor Modeling For CACD Purposes," IEEE Trans. on Electron Devices vol ED-32 pp 2415-2419, Nov 1985

Ebe54 Ebers J J and J L Moll "Large Signal Behavior of Junction Transistors," Proc. IRE vol 42 pp 1761-1772 Dec 1954

Elm48 Elmore W C , "The Transient Response of Damped Linear Networks with Particular Regard to Wideband Amplifiers " Journal of Applied Physics, vol 19 pp 55-63 Jan 1948

Eng69 Engl W L and J B Kioustelidis "Use of Padé Approximations in Device Modelling " Solid-State Electronics, vol 12, pp 239-246 1969

Eng83 Engl W L MEDUSA-User Manual, ITHE Aachen, Techn Hochschule Aachen, Aachen Germany 1983

Fos86 Fossum J G and S Veeraraghavan, "Partitioned Charge-Based Modeling of Bipolar Transistors for Non Quasi-Static Circuit Simulation," IEEE Electron Device Letters vol EDL 7 pp 652-654, Dec 1986

Ger78 Getreu I , Modeling the Bipolar Transistor, Elsevier Scientific Publishing Company New York pp 151-164 1978

Gia54 Giacoletto L J , "Study of P/N/P Alloy Junction Transistor From D C Through Medium Frequencies," RCA Rev. vol 15, p 506 1954

Gou82 Gough, R G "High Frequency Transistor Modeling for Circuit Simulation " IEEE J. Solid-State Circuits vol SC-17 no 4, pp 666-670 Aug 1982

Gre64 Grey, P E D DeWitt, A R Hoothroyd and J F Gibbons Physical Electronics and Circuit Models of Transistors, John Wiley and Sons, Inc , New York pp 214-222 1964

Gre69 Grey P E , and C L Searle Electronic Principles -Physics Models and Circuits, John Wiley and Sons New York pp 293-296, 336-341 1969

Gre84 Gray P R and R G Meyer Analysis and Design of Analog Integrated Circuits 2ed , John Wiley & Sons New York pp 363 367 1984

Gre84 Crebene A B Bipolar and MOS Analog Integrated Circuit Design John Wiley & Sons New York pp 220-223 1984

Gum68 Gummel H K "A Charge Control Transistor Model for Network Analysis Programs." Proc. IEEE (Letters) vol 57 p 751 Apr 1968

Gum70a Gummel H K "A Charge Control Relation for Bipolar Transistors." Bell Syst. Tech. J. vol 49 p 115, 1970

Gum70b Gummel H K, and H C Poon, "An integral charge control model of bipolar transistors," Bell Syst. Tech. J. vol 49 p 827, 1970

Ham63 Hamilton D J F A Lindholm, and J A Nerud, "Large Signal Models for Junction Transistors." Proc. of 1963 International Solid-State Circuits Conference, pp 56-58 Feb 1963

Ham64 Hamilton D J F A Lindholm, and J A Nerud, "Comparison of Large-Signal Models for Junction Transistors." Proc. IEEE vol 52 pp 239-248 Mar 1964

Ham71 Hamilton D J, F A Lindholm, and A N Mershak Principles and Applications of Semiconductor Device Modeling Holt, Rienhart and Winston Inc New York, 1971

Ham75 Hamilton D G and W C Howard Basic Integrated Circuit Engineering, McGraw-Hill New York pp 253 258, 1975

Hod83 Hodge D A, and H C Jackson Analysis and Design of Digital Integrated Circuits McGraw-Hill New York pp 228-241 1983

Joh59 Johnson E O and A Rose "Simple General Analysis of Amplifier Devices with Emitter, Control and Collector Functions," Proc. IRE vol 47, pp 407 418, March 1959

Jun86 Jung T A, "Use of Transients in Quasi Neutral Regions for Characterizing Solar Cells, Diodes, and Transistors," Ph D Dissertation, University of Florida 1986

Kak71 Kakihana S and F H Wang, "Simple CAD Technique to Develop High Frequency Transistors." IEEE Journal of Solid-State Circuits vol SC-6, no 4 pp 236-243 Aug 1971

Klo87 Klose H and A W Wieder "The Transient Integral Charge Control Relation: A Novel Formulation of the Currents in a Bipolar Transistor." IEEE Trans. on Electron Devices, vol ED-34, no 5, pp 1090-1099 May 1987

Lan72 Lange J and W N Carr, "An Application of Device Modeling to Microwave Power Transistors," IEEE Journal of Solid-State Circuits vol SC-7, no 1, pp 71-80 Feb 1972

Lin71a Lindholm F A "Integrated Circuit Transistor and Diode Models for Network Analysis Programs " IEEE Trans. on Circuit Theory CT-18 p 122 Jan 1971

Lin71b Lindholm F A "Unified Modeling of Field-Effect Devices " IEEE J. Solid State Circuits vol Sc-6 p 250 1971

Lin64 Lindholm F A and D J Hamilton "A Systematic Modeling Theory for Solid-State Devices " Solid-State Electronics vol 7 pp 771 783 Nov 1964

Lin85 Lindholm F A M K Chen B Wu, and T W Jung "Models Including Carrier Propagation For Bipolar Circuit Simulation," Proc. of SRC Topical Conference p 20 Apr 1986

Lin65 Lindmayer J , and C Y Wrigley Fundamentals of Semiconductor Devices Van Nostrand Co , Princeton NJ 1965

Lin58 Linvill J G "Lumped Models of Transistors and Diodes " Proc. IRE vol 46 pp 1141-1152 1958

Lin61 Linvill J G , and J F Gibbons, Transistors and Active Circuits McGraw-Hill New York Chap 9 1961

Leg71 Logan, J . "Statistical Circuit Design Characterization and Modelling for Statistical Design " Bell System Tech. J. vol 50 pp 1105-1147 1971

Lyn67 Lynn D K C S Meyer and D J Hamilton Analysis and Design of Integrated Circuits McGraw Hill, New York Chap 4 1967

Mac72 Macnee A B , and R J Telsky "High-Frequency Transistor Model for Circuit Design," IEEE J. of Solid-State Circuits, vol SC-6 pp 320-322, Aug 1972

Mer81 Mershek A H M A Shibib J C Fossum and F A. Lindholm, "Rigid Band Analysis of Heavily Doped Semiconductor Devices " IEEE Trans. on Electron Devices vol Ed-28 no 3, pp 293 297 Mar 1981

Mid59 Middlebrook R D , "A Modern Approach to Semiconductor and Vacuum Device Theory," Proc. IRE vol 106 Part B Suppl no 17, p 857 1959

Mil67 Millman J , and A Grebel, Microelectronics pp 567-569, McGraw-Hill New York 1967

Mil80 Milnes A C , Semiconductor Devices and Integrated Electronics pp 238-240 Van Nostrand Reinhold New York 1980

Mot83 Motz F R and W A Vincent "Automotive Electronics Designing Custom ICs for a Harsh Environment," Proc. of IEEE Custom Integrated Circuits Conference pp 392 398 1983

Neg77 Negel L W "SPICE2 A Computer Program to Simulate Semiconductor Circuits" Electronics Research Lab Univ California Berkeley Memo ERL M5D May 1975

Oh80 Oh S Y D E Ward and R W Dutton "Transient Analysis of MOS transistors" IEEE Trans. on Electron Devices vol ED-27, pp 1571-1578, Aug 1980

Oll72 Olline, R I and S J Retner "Computer Aided Design and Optimization of a Broad-band High Frequency Monolithic Amplifier," IEEE J. Solid State Circuits vol SC-7 pp 487 492 Dec 1972

Pau83 Paulos J J and D A Antoniadis, "Limitations of Quasi Static Capacitance Models for the MOS Transistor" IEEE Electron Device Letters vol EDL-4 no 7 pp 221 224 July 1983

Ros54 Rose, A , "La Photoconductivite," L'Onde Electrique, vol 34 pp 845 851 1954

Ros74 Ross, S , Introduction to Ordinary Differential Equations p 47, Xerox College Publishing, Toronto 1974

Sah64 Sah C T , "Characteristics of the Metal-Oxide Semiconductor Transistors," IEEE Trans. on Electron Devices vol ED 11 pp 324 345 July 1964

Sah70 Sah C T "The Equivalent Circuit Model in Solid-State Electronics III Conduction and Displacement Currents," Solid-State Electronics vol 13, pp. 1547 1575, 1970

Sah71 Sah C T "Equivalent Circuit Models in Semiconductor Transport for Thermal Optical Auger Impact and Tunneling Accumulation Generation trapping Process " Phy. Stat. Sol. (a)7 pp 541 559 16 Oct 1971

Sah73 Sah C T and F A Lindholm "Transport in Semiconductors with Low Scattering Rate and at High Frequencies," Solid State Electronics vol 16 pp 1447 1463, 1973

Sch83 Schutt Aine J E , "Determination of a Small-Signal Model for Ion-Implanted Microwave Transistors," IEEE Trans. on Electron Devices vol ED 30 no 7 pp 750 758 July 1983

Sei87 Seitchik J A , A Chatterjee, and P Yang, "An Accurate Bipolar Model for Large Signal Transient and ac Application " Proc. of IEDM pp 244 247 1987

She52 Shekel J "Matrix Representation of Transistor Circuits," Proc. IRE vol 40 pp 1493-1497 Nov 1952

Sha49 Shockley, W., "The Theory of p n Junctions in Semiconductors and p-n Junction Transistors " Bell System Technical Journal pp 435-489 1949

Sli86 SLICE Manual Rev 5.01 Harris Semiconductor Corporation Melbourne FL, July 1986

Tew73 Te Winkel J "Extended Charge Control Model for Bipolar Transistors " IEEE Trans. on Electron Devices vol ED-20 no 4 pp 389-394 Apr 1973

Tho58 Thomas D E and J L Moll "Junction Transistor Short-circuit Current Gain and Phase Determination," Proc. IRE vol 46, no 6 pp 1177-1184 June 1958

Tsi87 Tsividis Y. Operation and Modeling of the MOS Transistor McGraw Hill New York Chap. 7. 1987

Van76 Van der Ziel, A Solid State Physical Electronics III ed Chapter 15, Prentice-Hall New Jersey pp 335-339 1976

Van53 Van Roosbroeck "The Transport of Added Current Carriers in a Homogeneous Semiconductor." Phy. Rev. 91 282-289 1953

Vee86 Veeraraghavan, S J G Fossum and W R Eisenstadt "SPICE Simulation of SOI MOSFET Integrated Circuits " IEEE Trans. on Computer-Aided Design, vol CAD-5, no 4 pp 653-658 Oct 1986.

War78 Ward, D E and R W Dutton, "A Charge-Oriented Model for MOS Transistor Capacitances " IEEE J. Solid-State Circuits vol SC-13, pp 703-707, Oct. 1978.

Wu87 Wu B S H K Chen F A Lindholm and T W Jung "Carrier-Propagation-Delay (CPD) Models for Bipolar Circuit Simulation " Proc. of IEEE Custom Integrated Circuits Conference p 413 May 1987

Wu89 Wu B S and F A Lindholm "One Dimensional Non-Quasi Static Models for Arbitrarily and Heavily Doped Quasi-Neutral Layers in Bipolar Transistors." IEEE Trans. on Electron Devices, vol ED-36 no 4, pp 727-737 Apr 1989

Yua87 Yuan, J S and W R Eisenstadt "Flexible Bipolar Transistor Modeling Subroutines Implemented on the SLICE/SPICE Circuit Simulator." Technical Report No 87-001 Integrated Electronic Center, University of Florida Feb 1987

BIOGRAPHICAL SKETCH

Meng-Kai Chen was born in Taichung, Taiwan, Republic of China, on May 11, 1958. He received the B.S. degree from National Taiwan University in 1981 and the M.E. degree from the University of Florida in 1984, both in electrical engineering. Since 1983, he has been working toward the Ph.D. degree at the University of Florida. He is now working in the semiconductor division of Delco Electronics, GM/Hughes Corporation. His area of responsibility is modeling and characterization of semiconductor devices.

His general areas of interest include semiconductor device physics, technology, and device modeling for computer-aided design.

He is a member of Tau Beta Pi, Eta Kappa Nu, and Phi Kappa Phi.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy

Fredrik A Lindholm Chairman
Professor of Electrical Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Arnost Neugroschel
Professor of Electrical Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy

Jerry G Fossum
Professor of Electrical Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy

Timothy J. Anderson
Tim J Anderson
Professor of Chemical Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy

William R. Eisenstadt
William R Eisenstadt
Assistant Professor of Electrical Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy

Robert M Fox
Assistant Professor of Electrical Engineering

This dissertation was submitted to the Graduate Faculty of the College of Engineering and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy

August 1989

Dean, College of Engineering

Dean, Graduate School